

MATEMÁTICA

Curso de pós-graduação
“lato sensu”

PROBABILIDADE E ESTATÍSTICA

Marcos Santos de Oliveira
Daniela Carine Ramires de Oliveira

Universidade Aberta do Brasil
Núcleo de Educação a Distância
Universidade Federal de São João del-Rei



Pós-graduação “lato sensu”
Curso de Matemática

Probabilidade e Estatística

Marcos Santos de Oliveira

Daniela Carine Ramires de Oliveira

UFSJ
MEC / SEED / UAB
2009

O48p Oliveira, Marcos Santos de
Probabilidade e estatística / Marcos Santos de Oliveira ; Daniela
Carine Ramires de Oliveira . – São João del-Rei, MG : UFSJ, 2009.
87 p.

Apostila do curso de Pós-graduação “lato sensu” em Matemática.

1. Matemática – Estudo e ensino 2. Probabilidade 3. Estatística I.
Oliveira, Daniela Carine Ramires de I. Título.

CDU: 519.2



Reitor

Helvécio Luiz Reis

Coordenador UAB/NEAD/UFSJ

Heitor Antônio Gonçalves

Coordenadora do curso Educação Empreendedora

Rosângela Maria de Almeida Camarano Leal

Coordenador do curso Matemática

Carlos Alberto Raposo da Cunha

Coordenadores do curso Práticas de Letramento e Alfabetização

Gilberto Aparecido Damiano

Maria José Netto Andrade

Conselho Editorial

Adélia Conceição Diniz

Alessandro de Oliveira

Bernadete Oliviera Sidney Viana Dias

Betânia Maria Monteiro Guimarães

Frederico Ozanan Neves

Geraldo Tibúrcio de Almeida e Silva

Gilberto Aparecido Damiano

Guilherme Chaud Tizziotti

Ignácio César de Bulhões

Luiz Fernando de Carvalho

Maria do Carmo Santos Neta

Maria do Socorro Alencar Nunes Macedo

Maria José Netto Andrade

Marise Santana da Rocha

Rosângela Branca do Carmo

Terezinha Lombello Ferreira

Edição

Núcleo de Educação a Distância - NEAD-UFSJ

Conselho Editorial NEAD-UFSJ

Capa / Diagramação

Luciano Alexandre Pinto

Sumário

Pra começo de conversa...	05
Unidade I - Introdução à Estatística e Amostragem.	07
Aula 1 Noções de Estatística	09
1.1 Introdução	09
1.2 Classificação de Variáveis	12
1.3 Exercícios	15
Aula 2 Técnicas de Amostragem	16
2.1 Introdução	16
2.2 Amostragem Aleatória Simples (AAS)	17
2.3 Amostragem Sistemática (AS)	18
2.4 Amostragem Estratificada Proporcional (AEP)	19
2.5 Amostragem por Conglomerado (AC)	20
2.6 Exercícios	22
Unidade II - Exploração de Dados	23
Aula 1 Tabelas e Gráficos	25
1.1 Tabelas de Freqüências	25
1.2 Tabelas de Classes de Freqüências	26
1.3 Gráficos para as Variáveis Qualitativas	27
1.4 Gráficos para as Variáveis Quantitativas	29
1.5 Exercícios	33
Aula 2 Medidas de Posição e Dispersão	34
2.1 Mínimo, Máximo e Moda	34
2.2 Média e Mediana	35
2.3 Medidas Separatrizes: Quartis, Decis e Percentis	38
2.4 Amplitude, Variância e Desvio Padrão	39
2.5 Exercícios	41

Unidade III - Probabilidade	43
Aula 1 Introdução à Probabilidade	45
1.1 Processo ou Experimento Aleatório	45
1.2 Espaço Amostral e Evento	45
1.3 Definições de Probabilidade	48
1.4 Exercícios	52
Aula 2 Fundamentos de Probabilidades	55
2.1 Probabilidade Condicional	55
2.2 Independência de Eventos	56
2.3 Regra da Probabilidade Total	57
2.4 Teorema de Bayes	58
2.5 Exercícios	59
Unidade IV – Distribuições de Probabilidades	61
Aula 1 Variáveis Aleatórias Discretas	63
1.1 Introdução	63
1.2 Esperança Matemática e Variância	65
1.3 Distribuições de Probabilidades para Variáveis Aleatórias Discretas	66
1.3.1 Modelo Bernoulli	66
1.3.2 Modelo Binomial	67
1.3.3 Modelo Hipergeométrico	68
1.3.4 Modelo Poisson	69
1.4 Exercícios	71
Aula 2 Variáveis Aleatórias Contínuas	72
2.1 Introdução	72
2.2 Esperança Matemática e Variância	76
2.3 Distribuições de Probabilidades para Variáveis Aleatórias Contínuas	76
2.3.1 Modelo Uniforme	77
2.3.2 Modelo Normal	78
2.4 Exercícios	84
Referências	87

PARA COMEÇO DE CONVERSA...

A elaboração deste livro nasceu da vontade de produzir um material didático adequado ao Ensino a Distância (EAD) de Probabilidade e Estatística para o curso de Pós-Graduação *Lato Sensu* de Matemática da Universidade Federal de São João del-Rei (UFSJ). O livro foi escrito com o objetivo de apresentar, de forma resumida e didática, os conceitos mínimos que são considerados essenciais no estudo do tema. Isso não significa que o estudante deva se limitar ao estudo deste volume. Ao contrário, ele é o ponto de partida para busca de um conhecimento mais amplo e aprofundado sobre o assunto.

O livro está dividido em quatro unidades, contendo duas aulas cada uma. Ao final de cada aula incluímos exercícios que visam à aplicação imediata dos conceitos discutidos.

Esperamos que o(a) prezado(a) Estudante sinta o prazer de estudar este livro na mesma proporção que os autores sentiram ao elaborar cuidadosamente cada conteúdo apresentado. **Atenção!** Recomendamos insistentemente que você estude uma aula por semana. Faça todos os exercícios propostos antes de iniciar o estudo da aula seguinte e tire suas dúvidas com os tutores presenciais e a distância. Lembre-se de que o ensino a distância tem suas peculiaridades e que você é o principal responsável pelo seu sucesso no curso. Por isso, é necessário que você tenha disciplina, dedicação e empenho. Não deixe acumular matéria. Caso isso aconteça, aproveite os fins de semana para colocar a matéria em dia e finalizar cada unidade proposta.

Nós, professores-autores, bem como os tutores presenciais e os tutores a distância, estamos à sua disposição para atendê-lo(a) da melhor maneira possível.

Agradecemos à equipe do NEAD/UFSJ pelo apoio na produção deste material. Pedimos desde já desculpas pelos erros que serão eventualmente identificados neste livro. As críticas e sugestões de colegas e estudantes serão muito bem-vindas e auxiliarão a melhoria da próxima versão.

Os Autores

UNIDADE I

INTRODUÇÃO À ESTATÍSTICA E AMOSTRAGEM

Objetivos

Ao final desta unidade, você deverá ser capaz de

1. Identificar população e amostra.
2. Conceituar e classificar variáveis.
3. Aplicar diferentes técnicas de amostragem.
4. Diferenciar as técnicas de amostragem a partir de suas características.

Aula 1 – Noções de Estatística

1.1 Introdução

A palavra estatística é derivada da palavra latina *status* (que significa “estado”). Os primeiros usos da estatística envolviam compilação de dados e gráficos que descreviam vários aspectos de um estado ou país. Em 1662, John Graunt publicou informação estatística acerca de nascimentos e mortes. O trabalho de Graunt foi seguido por estudos sobre taxas de mortalidade e de doenças, tamanhos de população, renda e taxas de desemprego. As famílias, os governos e as empresas se apóiam fortemente nos dados estatísticos para orientação. Por exemplo, taxas de desemprego, taxas de inflação, índices do consumidor e taxas de nascimento e morte são cuidadosamente compiladas de modo regular, e os dados resultantes são usados para tomar decisões que afetam futuras contratações, níveis de produção e expansão para novos mercados. Assim, necessitamos entender os conceitos básicos da Estatística, bem como as suposições necessárias para o seu emprego de forma criteriosa, em cada tipo de problema a ser analisado.

O que é Estatística?

Podemos considerar que a Estatística é uma ciência que fornece um conjunto de técnicas que nos permitem, coletar, organizar, descrever, analisar e interpretar dados oriundos de estudos ou experimentos realizados em qualquer área do conhecimento. Estamos denominando por dados a um (ou mais) conjunto de valores, numéricos ou não. A aplicabilidade das técnicas a serem discutidas se dá nas mais variadas áreas das atividades humanas. Nesse sentido, o principal objetivo da Estatística é nos auxiliar a tomar decisões ou tirar conclusões em situações de incerteza, a partir de informações numéricas.

A Estatística pode ser dividida em três áreas, a saber:

- **Estatística Descritiva:** conjunto de técnicas destinadas a descrever e resumir os dados, a fim de tirarmos conclusões a respeito de características de interesse.

- **Probabilidade:** teoria matemática utilizada para se estudar a incerteza associada a fenômenos aleatórios.
- **Inferência Estatística:** denominação usualmente empregada ao estudo de técnicas que possibilitam a extrapolação, a um grande conjunto de dados (população), das informações e conclusões obtidas a partir de um subconjunto de valores (amostra).

Estudos complexos envolvendo o tratamento estatístico dos dados usualmente envolvem as três áreas mencionadas anteriormente. Para exemplificar tal procedimento, considere o esquema apresentado na Figura 1, a seguir:

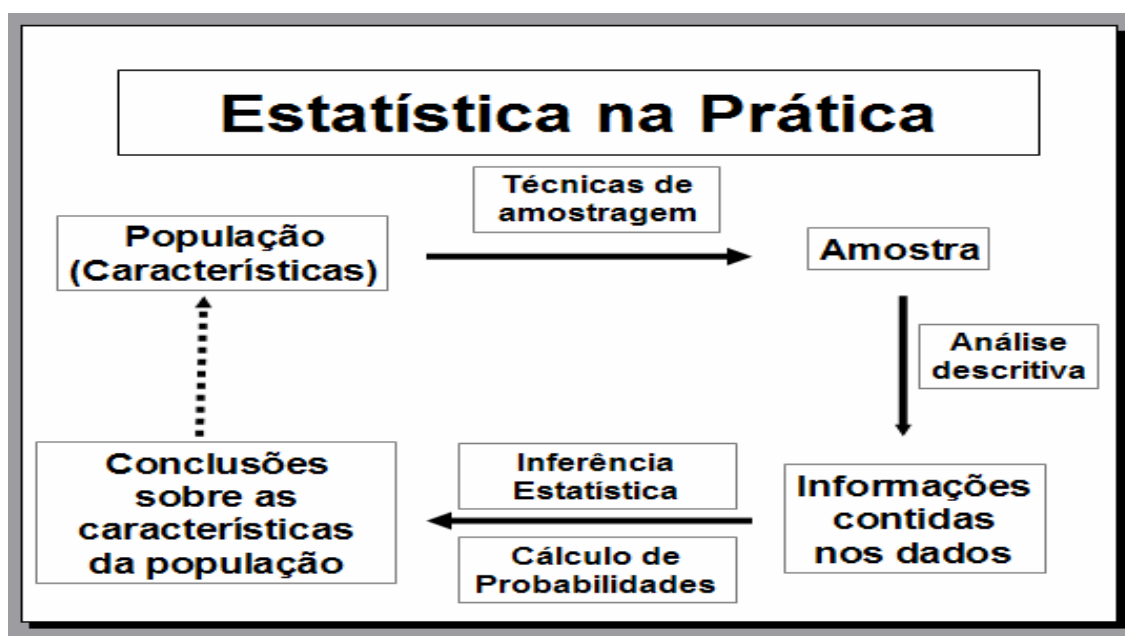


Figura 1. Estatística na prática.

A Figura 1 ilustra como a Estatística funciona na prática. Suponha, inicialmente, que estamos interessados em estudar algumas características em um grande conjunto de dados que denominaremos de **população**. Deve-se considerar que, na terminologia estatística, população refere-se não somente a uma coleção de indivíduos, mas ao alvo no qual reside nosso interesse. Assim, todos os clientes de um banco, todos os alunos de uma faculdade, todos os automóveis de uma determinada marca, ou mesmo todo o sangue no corpo de uma pessoa são

exemplos de possíveis populações. Algumas vezes podemos acessar todos os dados da população (nesse caso dizemos que o **censo** foi realizado), mas em muitas situações tal procedimento não pode ser realizado. Em geral, razões econômicas e éticas são as mais determinantes dessas situações. Para contornar esse fato, tomamos alguns elementos da população para formar um grupo a ser estudado. Esse subconjunto da população, em geral com dimensão sensivelmente menor, é denominado **amostra**.

A seleção de uma amostra pode ser feita de várias maneiras, dependendo, entre outros fatores, do grau de conhecimento que temos da população, da quantidade de recursos disponíveis, e assim por diante. Existem técnicas adequadas de amostragem que nos auxiliam na obtenção de um subconjunto de valores o mais parecido possível com a população que lhe dá origem. Algumas dessas técnicas serão vistas posteriormente.

Obtida uma amostra, o próximo passo é utilizar as técnicas de **Estatística Descritiva** para organizar e descrever os resultados contidos na amostra. A partir daí, podemos usar técnicas de **Inferência Estatística** para estimar quantidades desconhecidas, realizar extrapolação dos resultados e testar algumas hipóteses de interesse sobre a população. As técnicas de Inferência Estatística não fazem parte da ementa desta disciplina; entretanto, as mesmas serão vistas de forma detalhada na disciplina **Estatística Aplicada**.

Um Pouco da História da Ciência Estatística

A título de curiosidade, apresentamos um pouco da história da Ciência Estatística.

5000 a.C. Surgiram os primeiros registros egípcios de presos de guerra.

2000 a.C. Houve o primeiro censo Chinês.

695 Primeira utilização da média ponderada pelos árabes na contagem de moedas.

1303 Origem dos números combinatórios (Shihcieh Chu).

1654 Pierre de Fermat e Blaise Pascal, dois famosos matemáticos, estabelecem os Princípios do Cálculo das Probabilidades.

1763 Primeiras idéias das técnicas de Inferência Estatística (Thomas Bayes).

1930 Início das técnicas de Controle Estatístico de Qualidade nas indústrias.

1940 Invenção do Computador Eletrônico.

Maiores detalhes sobre a história da Estatística podem ser encontrados no site da Associação Brasileira de Estatística – ABE, no link

<http://www.redeabe.org.br/historia.htm>

1.2 Classificação de Variáveis

Qualquer característica associada a uma população é chamada de **variável**. Ela recebe esse nome porque ela “varia” de alguma forma. A idade de um indivíduo, o sexo ou o estado civil são possíveis exemplos de variáveis. Alguns conjuntos de dados consistem de números (tais como altura de 1,50 m a 2,15 m), enquanto outros são não-numéricos (tais como cor dos olhos: verde e castanho). Os termos *dados quantitativos* e *dados qualitativos* são em geral usados para distinguir entre esses dois tipos. Dessa forma, as variáveis podem ser classificadas como **Qualitativas** ou **Quantitativas**. Vejamos um exemplo.

Exemplo 1. A MD Indústria e Comércio, desejando melhorar o nível de seus funcionários, montou um curso experimental e indicou 25 funcionários para a primeira turma. Os dados estão dispostos na Tabela 1. Como havia dúvidas quanto à adoção de um único critério de avaliação, cada instrutor adotou seu próprio sistema de aferição.

De modo geral, para cada elemento investigado numa pesquisa, tem-se associado um (ou mais de um) resultado correspondendo à realização de uma característica (ou características). Por exemplo, considerando a variável *conceito em inglês*, para cada funcionário pode-se associar um dos resultados, *A, B, C ou D*.

Tabela 1. Informações sobre *seção, grau de instrução, números de filhos*, notas e conceitos nas disciplinas *redação, inglês, metodologia e política* de 25 empregados da MD Indústria e Comércio.

Func.	Seção	Grau de instrução	Nº de filhos	Redação	Inglês	Metodologia	Política
1	Pessoal	Ensino Médio	0	8,6	B	A	9,0
2	Pessoal	Fundamental	2	7,0	B	C	6,5
3	Pessoal	Ensino Médio	3	8,0	D	B	9,0
4	Pessoal	Ensino Médio	1	8,6	D	C	6,0
5	Pessoal	Superior	2	8,0	A	A	6,5
6	Pessoal	Superior	1	8,5	B	A	6,5
7	Pessoal	Fundamental	1	8,2	D	C	9,0
8	Técnica	Fundamental	2	7,5	B	C	6,0
9	Técnica	Superior	3	9,4	B	B	10,0
10	Técnica	Ensino Médio	4	7,9	B	C	9,0
11	Técnica	Fundamental	2	8,6	C	B	10,0
12	Técnica	Ensino Médio	3	8,3	D	B	6,5
13	Técnica	Superior	1	7,0	B	C	6,0
14	Técnica	Superior	1	8,6	A	B	10,0
15	Venda	Ensino Médio	0	8,6	C	B	10,0
16	Venda	Fundamental	1	9,5	A	A	9,0
17	Venda	Superior	0	6,3	D	C	10,0
18	Venda	Fundamental	0	7,6	C	C	6,0
19	Venda	Superior	3	6,8	D	C	6,0
20	Venda	Superior	2	7,5	C	B	6,0
21	Venda	Fundamental	1	7,7	D	B	6,5
22	Venda	Ensino Médio	2	8,7	C	A	6,0
23	Venda	Fundamental	1	7,3	C	C	9,0
24	Venda	Superior	0	8,5	A	A	6,5
25	Venda	Superior	1	7,0	B	A	9,0

Fonte: Adaptado de Bussab e Morettin (2006).

Algumas variáveis como *seção, grau de instrução, conceito em inglês e conceito em metodologia* apresentam como possíveis resultados uma qualidade (ou atributo) do indivíduo pesquisado. Logo, essas variáveis são chamadas de **variáveis qualitativas**. Dentre as variáveis qualitativas, ainda podemos fazer uma distinção entre dois tipos, a saber: *variável qualitativa nominal* ou *variável qualitativa ordinal*.

Uma variável é **qualitativa nominal** se não existe nenhuma ordenação nos possíveis resultados. Possíveis exemplos são *seção a que o funcionário pertence, sexo, raça* etc.

Uma variável é **qualitativa ordinal** se existe uma ordem natural nos seus resultados. Alguns exemplos são *grau de instrução, conceito em inglês, classe social* etc.

As variáveis *nota em redação, nota em política e número de filhos* apresentam como possíveis resultados números resultantes de uma contagem ou mensuração. Essas variáveis são chamadas de **variáveis quantitativas**. As variáveis quantitativas também podem sofrer uma classificação dicotômica: discreta ou contínua.

Uma variável é **quantitativa discreta** se os seus possíveis valores formam um conjunto finito ou infinito enumerável de números, e que resultam, freqüentemente, de uma contagem. Alguns exemplos são *números de filhos, números de carros na família* etc.

Uma variável é **quantitativa contínua** se os seus possíveis valores pertencem a um intervalo de números reais e que resultam de uma mensuração. Possíveis exemplos são *nota em redação e política, peso, altura* etc.

Para cada tipo de variável existem técnicas apropriadas para resumir as informações dos dados obtidos da amostra. Por exemplo, a utilização de uma tabela é um meio de descrever os dados de uma forma resumida. Veremos mais detalhes sobre tabelas e gráficos nas próximas seções.

Em algumas situações podemos atribuir valores numéricos às várias qualidades ou atributos de uma variável qualitativa e depois se proceder à análise como se esta fosse quantitativa, desde que o procedimento seja passível de interpretação.

Existe um tipo de variável qualitativa para a qual essa quantificação é muito útil: a chamada **variável dicotômica**. Para essa variável podem ocorrer somente duas realizações, usualmente chamadas de sucesso e fracasso. Exemplos de variáveis dicotômicas são *sexo, hábito de fumar (sim ou não)* etc.

1.3 EXERCÍCIOS

1. Para as situações descritas a seguir, identifique a população e a amostra correspondente e discuta a validade do processo de inferência estatística para cada um dos casos.

- a.** Uma amostra de sangue foi retirada de um paciente com suspeita de anemia.
- b.** Para verificar a audiência de um programa de TV, 563 indivíduos foram entrevistados por telefone com relação ao canal em que estavam sintonizados.
- c.** A fim de avaliar a intenção de voto para presidente dos brasileiros, 122 pessoas foram entrevistadas em Brasília.

2. Classifique cada uma das variáveis abaixo em qualitativa (nominal ou ordinal) ou quantitativa (discreta ou contínua).

- a.** Intenção de voto para presidente (possíveis respostas são os nomes dos candidatos, além de “não sei”).
- b.** Perda de peso de maratonistas na Corrida de São Silvestre, em quilos.
- c.** Intensidade da perda de peso de maratonistas na Corrida de São Silvestre (leve, moderada, forte).
- d.** Grau de satisfação da população brasileira com relação ao trabalho de seu presidente (valores de 0 a 5, com 0 indicando totalmente insatisfeito e 5 totalmente satisfeito).
- e.** Número de peças produzidas por uma máquina num dia de trabalho (500, 1000 etc).

Aula 2 – Técnicas de Amostragem

2.1 Introdução

A amostragem é naturalmente usada em nossa vida diária. Por exemplo, para verificar o tempero de um alimento em preparação, podemos provar (observar) uma pequena porção deste alimento. Nesse caso, estamos fazendo uma amostragem, ou seja, extraindo do **todo** (população) uma **parte** (amostra) com propósito de avaliarmos sobre a qualidade do tempero de todo o alimento.

Por que realizar amostragem?

Existem várias razões para o uso de amostragem em levantamento de grandes populações. Algumas delas, entre outras, são as seguintes:

- **Economia:** em geral, torna-se bem mais econômico o levantamento de somente uma parte da população.
- **Tempo:** numa pesquisa eleitoral, a três dias de uma eleição presidencial, não haveria tempo suficiente para pesquisar toda a população de eleitores do país.
- **Operacionalidade:** é mais fácil realizar operações de pequena escala. Um dos problemas típicos nos grandes censos é o controle dos entrevistadores.

Quando o uso de amostragem não é interessante?

- **População pequena:** não há necessidade de utilizar técnicas estatísticas, pois neste caso é aconselhável realizar o censo (análise de toda a população).
- **Característica de fácil mensuração:** talvez a população não seja tão pequena, mas a variável que se quer observar é de tão fácil mensuração que não compensa investir num plano de amostragem. Por exemplo, para verificar a porcentagem de funcionários favoráveis à mudança no horário de um turno de trabalho, podemos entrevistar toda a população no próprio local de trabalho. Esta atitude pode ser politicamente mais recomendável.

- **Necessidade de alta precisão:** a cada dez anos o IBGE¹ realiza um censo demográfico para estudar diversas características da população brasileira. Dentre estas características tem-se o número total de habitantes, uma informação fundamental para o planejamento do país. Dessa forma, o número de habitantes precisa ser avaliado com grande precisão e, por isso, se pesquisa toda a população.

2.2 Amostragem Aleatória Simples (AAS)

A técnica de amostragem aleatória é o método mais simples e um dos mais utilizados para a seleção de uma amostra. Para a seleção de uma AAS precisamos ter uma lista completa dos elementos da população. Este tipo de amostragem consiste em selecionar a amostra através de um sorteio. Sua principal característica está no fato de todos os elementos da população ter igual probabilidade de serem escolhidos.

- **Procedimento para o uso deste método**

1. Numerar todos os elementos da população (de 1 a N) e
2. Efetuar sucessivos sorteios até completar o tamanho da amostra (n).

Para realizar este sorteio, podemos utilizar urnas, tabelas de números aleatórios ou algum software que gere números aleatórios. A Tabela 2 foi construída usando-se o software Excel[®] (comando “aleatorio”).

Exemplo 2. Estamos interessados em estudar a qualidade da gasolina nos postos de uma determinada cidade. Essa cidade possui $N = 40$ postos. A empresa que estudará a qualidade pode investigar apenas uma amostra de $n = 4$ postos. Para selecionar uma amostra aleatória simples basta escolhermos uma posição de qualquer linha da tabela de números aleatórios e extrairmos conjuntos de dois algarismos (pois N, que é o tamanho da população, possui 2 casas decimais), até completarmos os 4 elementos da amostra. Se o número sorteado não

¹ IBGE - Instituto Brasileiro de Geografia e Estatística

existir, simplesmente não consideramos e prosseguimos o processo.

Escolhendo a primeira linha da tabela de números aleatórios, temos a seguinte amostra de 4 elementos:

$$AAS = \{16, 24, 18, 27, 25\}.$$

Tabela 2. Tabela de números aleatórios.

1	6	8	1	5	2	9	6	4	5	7	0	2	4	8	5	8	3	6	6	8	4	4	6	6	7	1	8	7	2	2	7	5	1	2	5	1	6	7	5
3	9	6	5	3	8	3	3	3	0	3	2	0	0	6	4	2	1	7	3	1	3	3	6	5	9	6	7	6	8	6	8	9	3	5	7	2	6	4	5
8	5	2	0	4	7	5	3	9	2	0	1	4	1	6	0	5	6	3	8	1	5	6	3	2	5	2	2	1	3	2	5	8	2	3	5	1	8	4	3
3	9	2	0	9	0	3	5	6	2	2	3	5	7	2	5	5	8	2	2	3	6	8	5	3	4	7	3	5	2	6	6	4	1	3	7	2	7	3	5
6	2	9	0	4	5	1	4	3	1	6	9	2	8	8	2	5	1	4	0	9	5	7	3	2	6	3	9	9	3	8	2	1	4	5	4	0	9	6	2
2	4	4	8	7	1	7	3	1	3	7	7	0	5	6	3	1	4	3	9	5	4	1	0	5	9	5	6	9	8	9	8	7	6	7	5	2	6	4	8
0	8	8	3	2	2	2	7	7	8	9	3	5	9	1	8	9	8	2	4	2	2	2	1	7	1	8	3	1	1	6	4	8	4	8	1	9	5	8	7
9	3	1	2	6	2	9	3	4	1	1	3	8	1	0	7	1	1	3	7	3	9	2	9	5	7	2	8	2	5	6	7	4	4	7	2	7	1	7	8
2	2	9	4	1	5	1	3	4	7	6	1	1	5	8	4	4	4	0	3	2	9	3	8	5	4	7	8	6	8	0	7	4	5	5	3	8	9	8	5
7	6	3	6	6	4	9	1	2	1	8	6	7	3	8	3	8	1	8	8	8	9	8	8	7	8	6	3	1	6	8	6	7	5	5	2	6	8	5	7
5	8	5	5	9	4	3	6	6	9	8	1	2	0	3	3	7	4	5	6	6	0	1	6	8	5	8	5	7	6	4	6	0	5	6	4	3	1	1	2
9	4	1	4	8	6	8	4	5	2	9	3	2	1	5	1	8	5	3	3	6	6	1	3	6	3	5	3	6	7	2	1	7	2	8	9	5	7	4	6
7	8	4	7	4	8	2	6	1	8	5	6	0	5	7	9	3	9	0	0	4	3	2	4	3	4	3	9	6	7	2	7	5	5	6	4	6	6	7	6
3	0	4	8	6	6	3	4	1	2	7	3	8	7	4	4	8	2	9	8	9	0	8	2	0	1	5	5	3	3	5	8	1	7	4	6	2	2	4	2
4	1	1	8	2	4	3	9	3	4	1	2	3	4	5	5	2	4	4	4	8	4	6	2	4	4	5	1	1	3	2	5	1	4	0	3	4	1	2	7

2.3 Amostragem Sistemática (AS)

É utilizada quando a população está naturalmente ordenada, como listas telefônicas, fichas de cadastramento e em sistemas de produções contínuos como produções de garrafas de cervejas etc.

▪ Procedimento para o uso deste método

1. Seja N o tamanho da população e n o tamanho amostral. Calcula-se o intervalo de amostragem $i = N/n$ (considera-se apenas a parte inteira do número i).
2. Sorteia-se, utilizando-se a tabela de números aleatórios, um número x entre 1 e i formando a amostra: $\{x, (x + i), (x + 2*i), \dots, (x + (n-1)*i)\}$.

Exemplo 3. Considerando uma turma com 49 alunos, retire uma amostra de tamanho 5 utilizando a técnica de amostragem sistemática.

Solução: Temos que $N = 49$ e $n = 5$. Logo,

1) $i = N/n = 49/5 = 9,8$. Considerando a parte inteira do número, temos que $i = 9$;

2) Sortear um número x entre 1 e $i = 9$ da tabela de números aleatórios que contenha um algarismo, pois i possui 1 casa decimal. Escolhendo a última linha, temos que o primeiro número que está entre 1 e 9 é 4. Logo, a amostra será composta dos seguintes elementos:

$$AS = \{4, 13, 22, 31, 40\}.$$

2.4 Amostragem Estratificada Proporcional (AEP)

A população é dividida em subgrupos, denominados estratos (por exemplo, por sexo, classe de renda, bairro etc.) e a AAS ou AS é utilizada na seleção de uma amostra de cada estrato. Esses estratos devem ser internamente mais **homogêneos** do que a população toda, com respeito às variáveis em estudo. Aqui, um conhecimento prévio sobre a população em estudo é fundamental.

A AEP tem as seguintes características:

- dentro de cada estrato há uma grande homogeneidade (pequena variabilidade);
- entre os estratos há uma grande heterogeneidade (grande variabilidade).

É comum os estratos terem tamanhos diferentes. Nesses casos, a proporcionalidade do tamanho da amostra de cada estrato da população deve ser mantida na amostra. Por exemplo, se um estrato corresponde a 20% do tamanho da população, ele também deve corresponder a 20% da amostra.

Exemplo 4. Com o objetivo de realizar uma pesquisa de opinião sobre a gestão atual da reitoria em uma determinada universidade, realizaremos um levantamento por amostragem. A população é composta por 70 professores, 80 servidores técnicos administrativos e 800 alunos,

que identificaremos da forma apresentada na Tabela 3.

Tabela 3. Listagem da população.

Professores	$P_{01} P_{02} \dots P_{70}$
Servidores	$S_{01} S_{02} \dots S_{80}$
Alunos	$A_{001} A_{002} \dots A_{800}$

Supondo que a opinião sobre a gestão atual da reitoria possa ser relativamente homogênea dentro de cada categoria, realizaremos uma amostragem estratificada proporcional por categoria, para obter uma amostra global de tamanho $n = 15$. A Tabela 4 mostra as relações de proporcionalidade.

Tabela 4. Relações de proporcionalidade.

Estrato	Proporção na população	Tamanho do subgrupo na amostra
Professores	$70/950 = 0,074$ (7,4 %)	$n_p = 15 \times 0,074 \approx 1$
Servidores	$80/950 = 0,084$ (8,4%)	$n_s = 15 \times 0,084 \approx 1$
Alunos	$800/950 = 0,842$ (84,2%)	$n_a = 15 \times 0,842 \approx 13$

Para selecionar aleatoriamente um professor, podemos usar a tabela de números aleatórios, tomando valores com dois algarismos. Usando a primeira linha, encontramos o seguinte professor selecionado: $\{P_{16}\}$. Para o servidor, usando a segunda linha da tabela, temos: $\{S_{39}\}$. Para os alunos, precisamos extrair números de três algarismos. Usando a terceira linha da tabela, temos: $\{A_{047}, A_{539}, A_{201}, A_{416}, A_{056}, A_{381}, A_{563}, A_{252}, A_{213}, A_{258}, A_{235}, A_{184}, A_{339}\}$. A amostra $\{P_{16}, S_{39}, A_{047}, A_{539}, A_{201}, A_{416}, A_{056}, A_{381}, A_{563}, A_{252}, A_{213}, A_{258}, A_{235}, A_{184}, A_{339}\}$ é uma amostra estratificada proporcional da comunidade da universidade. Cada indivíduo desta amostra deverá ser pesquisado para se obter a opinião em relação à gestão atual da reitoria.

2.5 Amostragem por Conglomerado (AC)

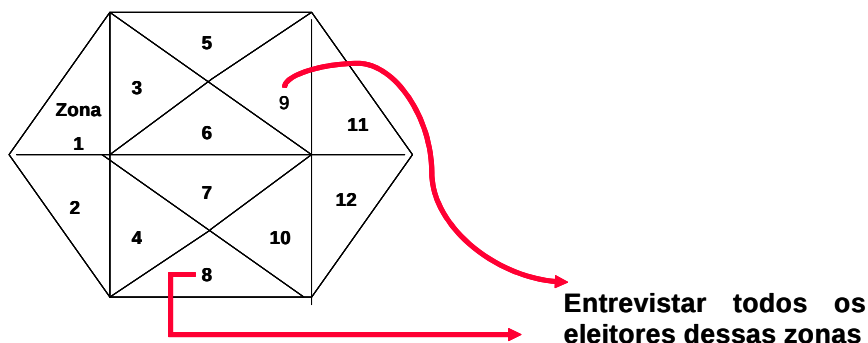
A população é dividida em subpopulações (conglomerados) distintas (quarteirões, residências,

famílias, bairros etc.). Alguns dos conglomerados são selecionados segundo a AAS, e todos os indivíduos nos conglomerados selecionados são observados. Em geral, é menos eficiente que a AAS ou AE, mas, por outro lado, é bem mais econômica. Tal procedimento amostral é adequado quando é possível dividir a população em um grande número de pequenas subpopulações.

A AC tem as seguintes características:

- dentro de cada conglomerado há uma grande heterogeneidade (grande variabilidade);
- entre os conglomerados há uma pequena variabilidade (grande homogeneidade).

Exemplo 5. Realização de uma pesquisa eleitoral em uma cidade com 12 zonas eleitorais. Usando a técnica de amostragem por conglomerados, podemos selecionar aleatoriamente 2 zonas eleitorais e, em seguida, entrevistar todos os eleitores dessas zonas selecionadas:



É fácil confundir amostragem estratificada com amostragem por conglomerado porque ambas envolvem a formação de subgrupos. A diferença é que a amostragem por conglomerado usa todos os membros de uma amostra de conglomerados, enquanto a amostragem estratificada usa uma amostra de membros de todos os estratos.

2.6 Exercícios

1. Refaça o Exemplo 4, considerando agora $n = 50$ indivíduos. Encontre todos os professores, funcionários e alunos que constituem a amostra estratificada proporcional.
2. Um administrador especialista em avaliar através de sistemas informatizados as ações da BOVESPA está interessado em fazer uma pesquisa nos preços das ações, para indicar aos seus clientes se hoje é um dia favorável a fazer investimentos. Ele sabe que existe $N = 500$ ações em venda. Como o tempo de estudo de cada ação é de aproximadamente 10 minutos, decidiu-se verificar apenas $n = 25$ ações. Utilizando-se as técnicas de amostragem aleatória simples e sistemática, quais ações serão selecionadas?
3. Um depósito em uma determinada empresa produtora de materiais eletrônicos possui $N = 100$ computadores que estão separados em duas qualidades: $N_1 = 40$ computadores Pentium 3 e $N_2 = 60$ computadores Pentium 4. O custo para verificar se cada computador está sob controle é muito alto. O administrador responsável disse que a empresa tem condições de verificar apenas $n = 12$ computadores. Utilizando-se a técnica de amostragem estratificada proporcional no primeiro estágio e a AAS no segundo estágio, quais computadores devem ser selecionados?
4. Uma cidade possui $N = 200$ zonas eleitorais. Uma empresa destinada a fazer uma pesquisa eleitoral vai selecionar aleatoriamente $n = 15$ zonas e entrevistar todos os elementos que estão dentro dessas zonas eleitorais, isto é, foi utilizada amostragem por conglomerado. Apresente quais serão as 15 zonas eleitorais amostradas.

UNIDADE II

EXPLORAÇÃO DE DADOS

Objetivos

Ao final desta unidade, você deverá ser capaz de

1. Organizar dados em tabelas de frequências e tabelas de classes de frequências.
2. Construir gráficos para variáveis qualitativas e quantitativas.
3. Calcular e interpretar medidas de posição.
4. Calcular e interpretar medidas de dispersão.

Aula 1 – Tabelas e Gráficos

1.1 Tabelas de Freqüências

Quando se estuda uma variável, o maior interesse do pesquisador é conhecer o comportamento dessa variável, analisando a ocorrência de seus possíveis resultados. Nesta seção veremos uma maneira de se dispor um conjunto de realizações, a fim de se ter uma idéia global sobre elas, ou seja, de sua distribuição.

Observando novamente a Tabela 1, especificamente a coluna que contém a variável *grau de instrução*, não conseguimos dizer rapidamente quantos funcionários possuem ensino fundamental, médio e superior. A Tabela 5 mostra uma maneira de representarmos mais resumidamente os dados da Tabela 1.

Exemplo 6. A Tabela 5 apresenta a distribuição de freqüências da variável *grau de instrução* dos dados da Tabela 1.

Tabela 5. Freqüências e porcentagens da variável *grau de instrução* para os 25 funcionários.

Grau de Instrução	Freqüência (n_i)	Proporção (f_i)	Porcentagem ($100 \times f_i$)
Fundamental	8	0,32	33,00
Ensino Médio	7	0,28	28,00
Superior	10	0,40	40,00
Total	25	1,00	100,00

Interpretação da Tabela 5. Nota-se que, dos 25 empregados, 33% tem nível fundamental, 28% nível médio e 40% nível superior.

Notação: Usaremos a notação n_i para indicar a **freqüência** (absoluta) de cada classificação ou categoria da variável. A notação $f_i = n_i/n$ para indicar a **proporção** (ou freqüência relativa) de

cada categoria, sendo o “**n**” o número total de observações.

As proporções (ou porcentagens) são muito úteis quando precisamos comparar resultados de duas pesquisas distintas. O próximo exemplo ilustra este fato.

Exemplo 7. Suponha que se queira comparar a variável *grau de instrução* dos empregados que fizeram o curso com a mesma variável para todos os empregados da Companhia MD. Digamos que a empresa tenha 2000 empregados e que a distribuição de frequências seja a da Tabela 6.

Tabela 6. Distribuição de frequências dos 2000 empregados segundo o *grau de instrução*.

Grau de Instrução	Frequência (n_i)	Proporção (f_i)	Porcentagem ($100 \times f_i$)
Fundamental	650	0,325	32,50
Ensino Médio	500	0,250	25,00
Superior	850	0,425	42,50
Total	2000	1,000	100,00

Comparação entre a Tabela 5 e a Tabela 6. Não podemos comparar diretamente as colunas das frequências (n_i) das duas tabelas, pois os totais de empregados são diferentes nos dois casos ($n = 25$ e $n = 2000$). Mas as colunas da proporção e da porcentagem são comparáveis, pois reduzimos a um mesmo total. Nesse caso, podemos dizer que a distribuição da variável *grau de instrução* dos funcionários que fizeram o curso não se diferencia da distribuição dessa mesma variável para todos os funcionários da Empresa MD.

1.2 Tabelas de Classes de Frequências

A construção de tabelas de frequências para variáveis quantitativas necessita de certo cuidado. Por exemplo, a construção da tabela de frequências para a variável *nota em redação* da Tabela 1, usando o mesmo procedimento de tabelas de frequências, não resumirá as 25 observações num grupo menor.

Solução: Agrupar os dados por faixas de notas. Assim, construímos a chamada tabela de classes de frequências.

Exemplo 8. A Tabela 7 fornece a distribuição de frequências das *notas em redação* dos 25 funcionários da Companhia MD por faixas de notas.

Tabela 7. Frequências e porcentagens das *notas em redação*.

Classe de notas	Frequência	Porcentagem
6 - 7	2	8
7 - 8	9	36
8 - 9	12	48
9 - 10	2	8
Total	25	100

Procedendo-se desse modo, ao resumir os dados referentes a uma variável quantitativa, perde-se alguma informação. Por exemplo, não sabemos quais são as doze notas da classe de 8 a 9, a não ser que investiguemos a tabela original. Sem perda de muita precisão, poderíamos supor que todas as doze notas daquela classe fossem iguais ao ponto médio da referida classe, isto é, 8,5.

A escolha dos intervalos é arbitrária. A familiaridade do pesquisador com os dados é que lhe indicará quantas e quais classes (intervalos) devem ser usadas. Entretanto, deve-se observar que, com um número pequeno de classes, perde-se informação, e com um número grande de classes, o objetivo de resumir os dados fica prejudicado. Normalmente, sugere-se o uso de 4 a 8 classes com a mesma amplitude.

1.3 Gráficos para Variáveis Qualitativas

A representação gráfica da distribuição de uma variável tem a vantagem de, rápida e concisamente, informar sobre sua variabilidade. Existem vários tipos de gráficos para as

variáveis qualitativas. Aqui serão ilustrados os dois mais simples e freqüentemente utilizados: gráficos de barras e de composição em setores (“pizza”).

Gráfico de barras

O gráfico de barras consiste em construir retângulos ou barras, em que uma das dimensões é proporcional à magnitude a ser representada (n_i), sendo a outra arbitrária, porém igual para todas as barras. Essas barras são dispostas paralelamente uma às outras, horizontalmente ou verticalmente. No exemplo a seguir temos o gráfico de barras (verticais) para a variável *grau de instrução* da Tabela 6.

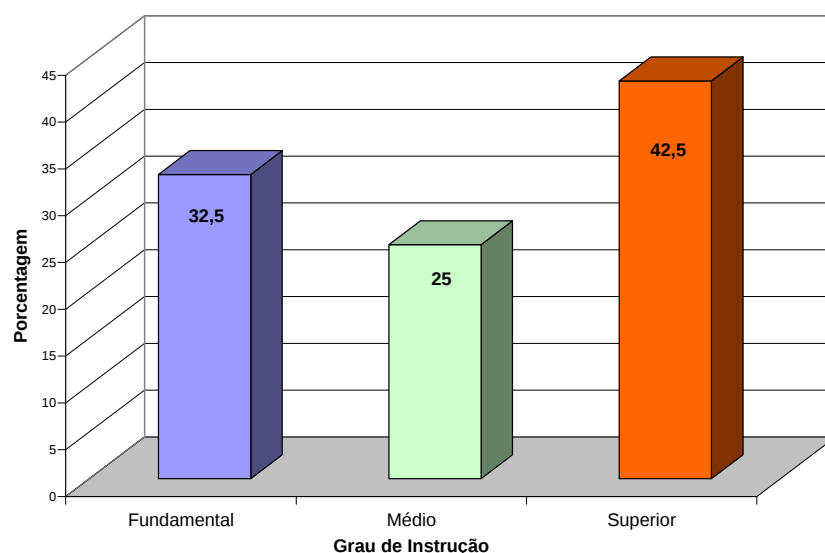


Figura 2. Gráfico de barras para a variável *grau de instrução*.

Gráfico de composição em setores (“pizza”)

O gráfico de composição em setores (“pizza”) destina-se a representar a composição, usualmente em porcentagem, de partes de um todo. Consiste num círculo de raio arbitrário, representando o todo, dividido em setores, que correspondem às partes de maneira proporcional. A Figura 3 ilustra esse gráfico para a variável *grau de instrução*.

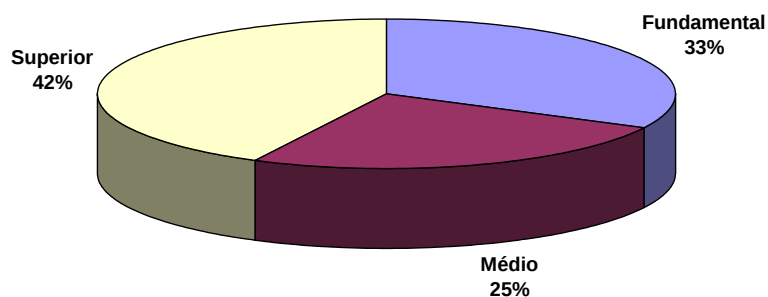


Figura 3. Gráfico em setores para a variável *grau de instrução*.

1.4 Gráficos para Variáveis Quantitativas

Para variáveis quantitativas podemos considerar uma variedade maior de representações gráficas.

Gráfico de barras

O gráfico de barras para as variáveis quantitativas é construído da mesma forma que o das variáveis qualitativas. Como ilustração, considere a variável *número de filhos* dos 25 empregados da Companhia MD. A Tabela 8 apresenta esses dados.

Tabela 8. Freqüências e porcentagens da variável *número de filhos*.

Nº de Filhos	Freqüência (n_i)	Porcentagem ($100 \times f_i$)
0	5	20
1	9	36
2	6	24
3	4	16
4	1	4
Total	25	100

A Figura 4 ilustra o gráfico de barras.

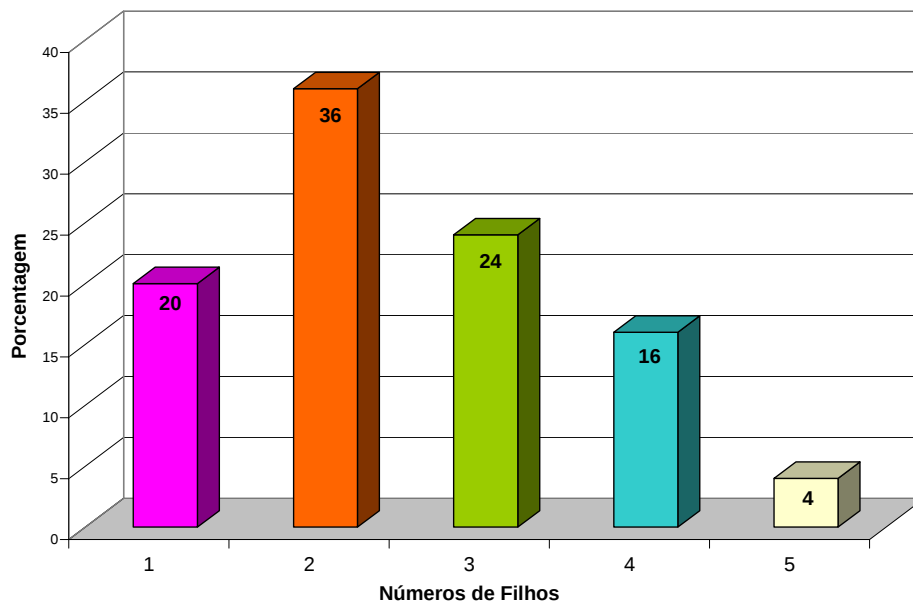


Figura 4. Gráfico de barra para a variável *número de filhos*.

Gráfico de pontos (Dot-Plot)

Quando os dados consistem em um pequeno conjunto de números, estes podem ser representados traçando-se uma reta com uma escala que abranja todas as mensurações observadas e grafando-se as respectivas freqüências como pontos acima da reta. Por esse motivo, é também conhecido como gráfico de pontos.

Exemplo 9. Considere a variável *tempo*, em segundos, entre carros que passam por um cruzamento, viajando na mesma direção. As 14 medições realizadas foram

6,0 3,0 5,0 6,0 4,0 3,0 5,0 4,0 6,0 3,0 4,0 5,0 2,0 11

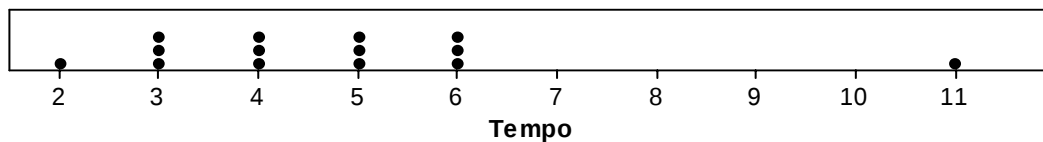


Figura 5. Gráfico de pontos para a variável *tempo*.

Histograma

O histograma consiste em retângulos contíguos com base nas faixas de valores da variável e com área igual à frequência relativa (f_i) da respectiva faixa. Desta forma, a altura de cada retângulo é denominada densidade de frequência definida pelo quociente da área pela amplitude da faixa, ou seja, f_i/a_i , com a_i indicando a amplitude da i -ésima classe. Com essa convenção, a área total do histograma será 1 (um).

Exemplo 10. Considerando a variável *nota em redação* dos 25 funcionários da Companhia MD, dispostos na Tabela 7. O histograma correspondente é apresentado na Figura 6.

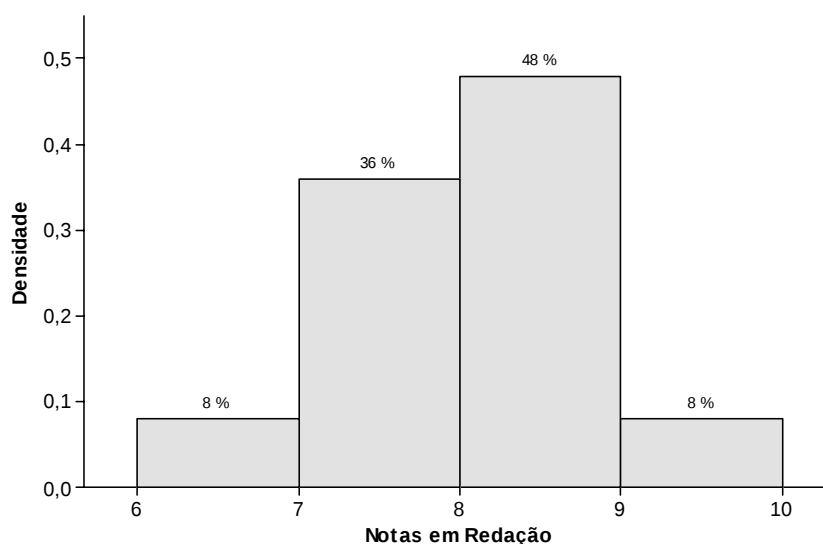


Figura 6. Histograma das *notas em redação*.

Gráfico de linhas

É um gráfico muito importante utilizado para representar observações feitas ao longo do tempo, em intervalos iguais ou não. Tais conjuntos de dados constituem as chamadas séries históricas ou séries temporais. Traduzem o comportamento de um fenômeno em certo intervalo de tempo.

Exemplo 11. Considere a dívida externa do Brasil (em milhões de dólares) no período de 1956 a 2006, apresentados na Tabela 9.

Tabela 9. Dívida externa do Brasil de 1956 a 2006, em milhões de dólares.

Ano	Dívida	Ano	Dívida	Ano	Dívida	Ano	Dívida
1956	2736	1969	4635	1982	85487	1995	159256
1957	2491	1970	6240	1983	93745	1996	179935
1958	2870	1971	8284	1984	102127	1997	199998
1959	3160	1972	11464	1985	105171	1998	241644
1960	3738	1973	14857	1986	111203	1999	241468
1961	3291	1974	20032	1987	121188	2000	236156
1962	3533	1975	25115	1988	113511	2001	226067
1963	3612	1976	32145	1989	115506	2002	227689
1964	3294	1977	37951	1990	123439	2003	235414
1965	3823	1978	52187	1991	123910	2004	220182
1966	3771	1979	55803	1992	135949	2005	187987
1967	3440	1980	64259	1993	145726	2006	191999
1968	4092	1981	73963	1994	148295		

Fonte: IPEADATA

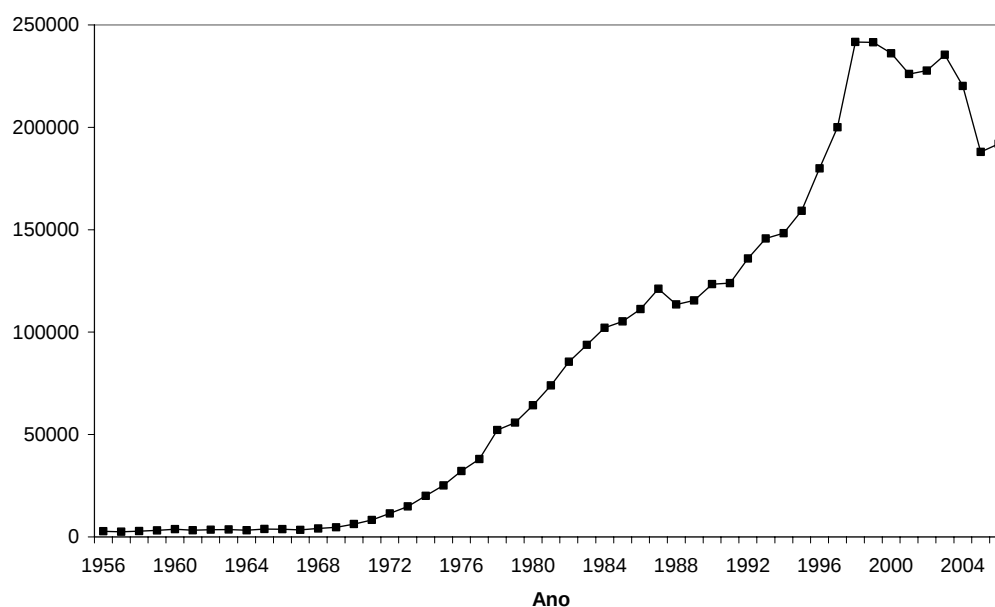


Figura 7. Gráfico de linhas da dívida externa do Brasil.

1.5 Exercícios

1. Os dados a seguir referem-se aos conceitos obtidos de 60 alunos, na disciplina de Estatística, de uma turma da UFSJ.

Tabela 10. Dados Brutos da disciplina de Estatística de uma turma da UFSJ.

R: Ruim					M: Médio					B: Bom					O: Ótimo				
M	R	M	M	M	R	B	B	M	M	R	B	M	M	M	M	R	B	B	R
B	M	R	M	B	M	R	M	R	M	B	M	R	M	R	M	B	M	B	M
B	B	B	B	O	M	M	M	M	M	B	B	B	B	B	B	B	O	B	O

a. Organize os dados da Tabela 10 em uma tabela de frequências contendo título, frequência absoluta, frequência relativa, porcentagens e uma interpretação.

b. Represente os dados da tabela obtido em a. através do gráfico de composição de setores.

2. A partir da Tabela 1, construa

a. a distribuição de frequências da variável *conceito em metodologia*, com as frequências absoluta e relativa, as porcentagens, dê um título e interprete;

b. uma tabela de classes de frequências para a variável *nota em política*, com as frequências absoluta e relativa, as porcentagens, dê um título e interprete;

c. Construa o gráfico de barras para a tabela montada no item a;

d. Faça o histograma utilizando a tabela de classes obtida do item b.

3. Faça o gráfico de linhas para os dados fornecidos na sua conta de luz durante o último ano, isto é, no eixo x coloque os meses e no eixo y coloque o consumo em kwh.

Aula 2 – Medidas de Posição e Dispersão

2.1 Mínimo, Máximo e Moda

O **mínimo** é a menor observação do conjunto de dados, enquanto que o **máximo** é a maior observação.

Exemplo 12. Considere o seguinte conjunto de dados: 4, 5, 4, 6, 5, 8, 4. Nesse caso, o mínimo é 4 e o máximo é 8.

A **moda** é o valor ou atributo que ocorre com maior frequência.

Exemplo 13. Considere os seguintes bancos de dados:

- a) 2, 5, 2, 7, 8 Neste caso a moda = 2.
- b) 3, 4, 2, 2, 4, 5 As modas são 2 e 4. Dizemos que o conjunto é bimodal.
- c) 1, 2, 3, 4, 5 O conjunto não apresenta moda, sendo chamado de conjunto amodal.

Podemos calcular o mínimo, máximo e moda se os dados estão agrupados em tabelas de frequências. Considere o próximo exemplo.

Exemplo 14. Uma empresa de segurança deseja estudar qual o número de ligações a cobrar mais frequentes que são recebidas em um determinado bairro de classe alta da cidade de São Paulo no mês de março. Foram selecionadas 30 residências e observado o *número de ligações a cobrar* em cada residência. O resultado se encontra na Tabela 11.

Tabela 11. Distribuição de frequência do *número de ligações a cobrar*.

Número de ligações a cobrar	Número de residências (n_i)
0	2
1	5
2	15
3	8
Total	30

A moda é 2 ligações a cobrar, pois foi o número que ocorreu com maior frequência. O valor mínimo foi zero e o valor máximo da variável foi 3.

2.2 Média e Mediana

A mais importante medida de posição é a média aritmética. Esse conceito já é, sem dúvida, familiar ao Leitor, quando fala, por exemplo, da altura média de um grupo de alunos ou da nota média da sala em determinada prova.

A **média aritmética** é a soma das observações divididas pelo número delas. De forma mais formal, considere n observações de um conjunto de dados representados por $x_1, x_2, \dots, x_n$. A média deste conjunto é obtida pela soma das n observações divididas por n , ou seja,

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n} \quad (4.1)$$

Exemplo 15. Considere o seguinte conjunto de notas: 2, 5, 3, 7, 8. A média das notas é

$$\bar{x} = \frac{2 + 5 + 3 + 7 + 8}{5} = \frac{25}{5} = 5$$

Podemos adaptar a fórmula (4.1) para o caso de dados agrupados em tabelas de frequência. Neste caso, a média é calculada levando-se em conta as frequências de cada valor da variável, da seguinte forma:

$$\bar{x} = \frac{\sum_{i=1}^v x_i n_i}{n} \quad (4.2)$$

onde v é a quantidade de resultados que a variável contém e n_i é a respectiva frequência da i -ésima classe. Assim, para o Exemplo 14, temos

$$\bar{x} = \frac{\sum_{i=1}^n x_i n_i}{n} = \frac{0 \times 2 + 1 \times 5 + 2 \times 15 + 3 \times 8}{30} = 1,9\bar{6} \cong 2.$$

Portanto, o número médio de ligações a cobrar recebido em um determinado bairro de classe alta da cidade de São Paulo no mês de março é 2.

A **mediana** é o valor que ocupa a posição central da série de observações, quando estão ordenadas em ordem crescente.

Assim, se as cinco observações de uma variável forem 3, 4, 7, 8 e 8, a mediana é o valor 7, correspondente à terceira observação. Quando o número de observações for par, usa-se como mediana a média aritmética das duas observações centrais. Acrescendo-se o valor 9 à série acima, a mediana será $(7 + 8)/2 = 7,5$.

Vamos formalizar o conceito da mediana. Considere que $x_1, x_2, \dots, x_n$ são os n valores (distintos ou não) da variável X . Considerando as observações ordenadas em ordem crescente, podemos denotar a menor observação por $x_{(1)}$, a segunda por $x_{(2)}$, e assim por diante, obtendo-se

$$x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n-1)} \leq x_{(n)} \quad (4.3)$$

Por exemplo, se $x_1 = 3, x_2 = -2, x_3 = 6, x_4 = 1$ e $x_5 = 3$, então $-2 \leq 1 \leq 3 \leq 3 \leq 6$, de modo que $x_{(1)} = -2, x_{(2)} = 1, x_{(3)} = 3, x_{(4)} = 3$ e $x_{(5)} = 6$.

As observações ordenadas como em (4.3) são chamadas estatísticas de ordem.

Com essa notação, a mediana da variável X pode ser definida como

$$\text{med}(x) = \begin{cases} x_{\left(\frac{n+1}{2}\right)} & \text{se } n \text{ é ímpar} \\ \frac{x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}}{2} & \text{se } n \text{ é par} \end{cases}$$

Nota: A mediana depende da posição e não dos valores dos elementos na série ordenada. Essa é uma diferença marcante entre mediana e média, pois a média se deixa influenciar, e muito, pelos valores extremos. Vejamos:

Na série: 5, 7, 10, 13, **15** Média = 10 e Mediana = 10;

Na série: 5, 7, 10, 13, **65** Média = 20 e Mediana = 10,

isto é, a média do segundo conjunto de valores é maior do que a do primeiro, por influência dos valores extremos, ao passo que a mediana permanece a mesma.

Quando os dados estão agrupados em tabelas de freqüências, o método mais prático para calcular a mediana é adicionar uma coluna à tabela contendo a freqüência acumulada. Vejamos um exemplo.

Exemplo 16. Considere novamente o Exemplo 14 da empresa de segurança que desejava estudar qual o número de ligações a cobrar mais freqüentes recebidas em um determinado bairro de classe alta da cidade de São Paulo no mês de março. Vamos introduzir uma nova coluna na tabela dos dados referente à freqüência acumulada, que é obtida acumulando-se as freqüências absolutas (n_i). No caso em particular teremos $F_1 = n_1$, $F_2 = n_1 + n_2$, $F_3 = n_1 + n_2 + n_3$ e finalmente, $F_4 = n_1 + n_2 + n_3 + n_4 = n$.

Como o rol é par, pois $n = 30$, a mediana será a média dos valores que estão nas posições 15ª e 16ª. Ambos os valores que estão nestas posições são 2 ligações a cobrar recebida por residência, pois F_3 é a primeira freqüência acumulada que contém os elementos da 15ª e 16ª posições.

Tabela 12. Frequência absoluta e acumulada do número de ligações a cobrar.

Número de ligações a cobrar	Número de Residências (n_i)	Freq. Acumulada (F_i)
0	2	2
1	5	7
2	15	22
3	8	30
Total	30	

2.3 Medidas Separatrizes: Quartis, Decis e Percentis

Além das medidas de posição que estudamos, há outras que, consideradas isoladamente, não são medidas de tendência central, mas estão ligadas à mediana relativamente à sua característica de separar a série em duas partes que apresentam o mesmo número de valores. Essas medidas - os quartis, os decis e os percentis - são, juntamente com a mediana, conhecidas pelo nome de separatrizes.

Denominamos **quartis** os valores de uma série que a dividem em quatro partes iguais. Portanto, precisamos de 3 quartis (Q_1 , Q_2 e Q_3) para dividir a série em quatro partes iguais. Note que o quartil 2 (Q_2) é por definição a própria mediana da série.

O método mais prático para calcular os quartis é utilizar o princípio do cálculo da mediana para os 3 quartis. Na realidade serão calculadas “3 medianas” em uma mesma série.

Exemplo 17. Considere a seguinte série de dados: 5, 2, 6, 9, 10, 13, 15. Ordenando a série, temos: 2, 5, 6, 9, 10, 13, 15. O valor que divide a série acima em duas partes iguais é 9. Logo a mediana é $9 = Q_2$. Temos agora {2, 5, 6} e {10, 13, 15} como sendo os dois grupos de valores iguais proporcionados pela mediana. Para o cálculo do quartil 1 (Q_1) e quartil 3 (Q_3) basta calcular as medianas de cada um desses grupos. Assim, em {2, 5, 6}, a mediana é $5 = Q_1$. Em {10, 13, 15} a mediana é $13 = Q_3$.

Seguindo o mesmo princípio dos quartis (que divide em quatro partes a série de dados) e levando em conta o aumento do número de informações disponíveis, podemos dividir a série de dados em 10 partes ou 100 partes. Quando dividimos em 10 partes, obtemos os **decis** ($D_1, D_2, \dots, D_9$) e em 100 partes obtemos os **percentis** ($P_1, P_2, \dots, P_{99}$).

Como ilustração, o decil D_6 representa o valor que deixa 60% das informações a sua esquerda e, conseqüentemente, 40% a sua direita. De forma análoga, o percentil P_{74} representa o valor que deixa 74% das observações a sua esquerda e 26% a sua direita.

2.4 Amplitude, Variância e Desvio Padrão

O resumo de um conjunto de dados por uma única medida representativa de posição central esconde toda a informação sobre a variabilidade do conjunto de observações. Começemos com um exemplo de motivação para ilustrar a importância da utilidade das medidas de dispersão, também conhecidas como medidas de variabilidade.

Exemplo 18. Para preencher uma única vaga existente em uma empresa, 50 candidatos foram submetidos a 6 provas de mesma importância sobre conhecimentos específicos de interesse da empresa. Três destes candidatos destacaram-se com as notas descritas na Tabela 13.

Tabela 13. Distribuição das notas.

Candidatos	Provas					
	1	2	3	4	5	6
A	7,0	7,5	8,0	8,0	8,5	9,0
B	6,0	7,0	8,0	8,0	9,0	10,0
C	7,5	8,0	8,0	8,0	8,0	8,5

Fonte: Dados hipotéticos

Que candidato escolher? Por um critério inicial poderia ser escolhido aquele com a maior média, mas todos têm mesma média, ou seja, 8. De modo análogo, nem adianta pensar em

moda ou mediana, pois também essas medidas são iguais a 8, para todos os candidatos.

Uma possível solução seria adotar um segundo critério: escolher o candidato que apresentou notas mais homogêneas, isto é, aquele que apresentou menor dispersão das notas. Poderíamos inicialmente calcular a **amplitude**, que é definida pelo intervalo entre o valor máximo e o valor mínimo da série de dados, ou seja, $A = \text{máx} - \text{min}$. Assim, teríamos as seguintes amplitudes: 2, 4 e 1, respectivamente para os candidatos A, B e C. Apesar de fácil de calcular, a amplitude tem a desvantagem de levar em conta apenas dois valores, desprezando todos os outros.

Uma medida de dispersão mais rica é obtida quando consideramos a soma dos quadrados dos desvios em relação à média. Essa medida é chamada de variância, sendo denotada por s^2 e definida por

$$s^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2}{n-1} = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} \quad (4.4)$$

A variância mede a dispersão dos dados em torno de sua média.

A raiz quadrada positiva da variância é chamada de desvio padrão (representado por s):

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} \quad (4.5)$$

Note que a unidade de medida do desvio padrão é a mesma dos dados originais, sendo assim interpretável, enquanto que a variância fornece uma unidade de medida elevada ao quadrado. O cálculo do desvio padrão exige o cálculo da variância.

Exemplo 19. A variância e o desvio padrão para o candidato A do Exemplo 18 fica

$$s_A^2 = \frac{(7-8)^2 + (7,5-8)^2 + (8-8)^2 + (8-8)^2 + (8,5-8)^2 + (9-8)^2}{6-1} = \frac{2,5}{5} = 0,5$$

$$s_A = \sqrt{0,5} \cong 0,7$$

De forma análoga podemos encontrar a variância e o desvio padrão para os candidatos B e C,

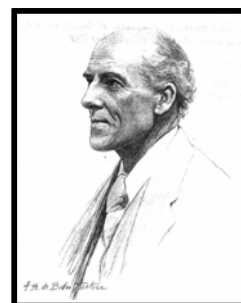
dados respectivamente por $s_B^2 = 2$ ($s_B \cong 1,4$) e $s_C^2 = 0,1$ ($s_C \cong 0,3$).

Podemos calcular a variância através da seguinte fórmula alternativa:

$$s^2 = \frac{1}{n-1} \left[\left(\sum_{i=1}^n x_i^2 \right) - n(\bar{x})^2 \right] \quad (4.6)$$

A fórmula (4.6) é obtida através de algumas manipulações algébricas na fórmula (4.4). Esta tem a facilidade de apenas necessitar da informação da média ($\bar{x}$) e da soma dos valores ao quadrado da variável $\left(\sum_{i=1}^n x_i^2 \right)$.

Karl Pearson



Um pouco de história

A primeira utilização do termo desvio padrão ocorreu em 1894, sendo devido Karl Pearson.

2.5 Exercícios

1. Os tempos de sobrevivência (em meses) de um tipo de bateria estão listados a seguir.

5, 21, 21, 23, 23, 25, 27, 29, 30, 31, 32, 32, 32, 34, 35, 36, 38, 38, 38, 42, 43, 44, 60.

a. Calcule a média e mediana. Comente os resultados.

b. Calcule o valor mínimo, Q_1 , Q_2 , Q_3 e máximo. Interprete estas 5 estatísticas.

c. Calcule a variância e desvio padrão. Comente.

2. Considere o seguinte conjunto de dados: 2, 3, 5, 7, 10. Utilize a formula alternativa para calcular a variância, sabendo que a média é 5,4.

3. Um órgão do governo do estado está interessado em determinar padrões sobre o investimento em educação, por habitante, realizado pelas prefeituras. De um levantamento de dez cidades, foram obtidos os valores (codificados) da tabela abaixo:

Cidade	A	B	C	D	E	F	G	H	I	J
Investimento	20	16	14	7	19	15	14	16	19	18

- Calcule a média das observações.
- Receberão um programa especial as cidades com valores de investimento inferiores à média menos duas vezes o desvio padrão. Alguma cidade receberá o programa?
- Será considerada como investimento básico a média das observações compreendidas entre a média original menos dois desvios padrão e a média original mais dois desvios padrão. Calcule o investimento básico e compare com a média obtida no item **a**. Justifique a diferença encontrada.

UNIDADE III

PROBABILIDADE

Objetivos

Ao final desta unidade, você deverá ser capaz de

1. Relacionar experimentos aleatórios com espaços amostrais.
2. Construir novos eventos a partir das operações elementares de eventos.
3. Calcular probabilidade a partir de eventos condicionais.
4. Calcular probabilidade a partir de eventos independentes.

Aula 1 – Introdução à Probabilidade

1.1 Processo ou Experimento Aleatório

Qualquer fenômeno que gere resultado incerto ou casual é chamado de **processo** ou **experimento aleatório**.

Exemplo 20. Os quatro itens a seguir ilustram experimentos aleatórios, pois não sabemos, com certeza, o possível resultado que ocorrerá em cada um.

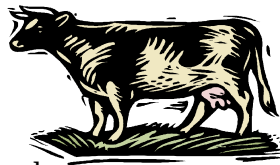
a. Jogar uma moeda duas vezes e observar a sequência obtida de caras e coroas.



b. Jogar um dado e observar o número mostrado na face superior.



c. Observar o peso de animais.



d. Observar o número de filhos de um casal.



1.2 Espaço Amostral e Evento

Espaço amostral (Ω) é o conjunto de todos os resultados possíveis de um experimento aleatório.

Todo experimento aleatório tem associado um espaço amostral. O Exemplo 21 ilustra esse fato.

Exemplo 21. Experimentos aleatórios e seus respectivos espaços amostrais.

Experimento aleatório	Espaço amostral
a. Jogar um dado e observar o resultado	$\Omega = \{1, 2, 3, 4, 5, 6\}$
b. Lançar uma moeda duas vezes e observar as faces obtidas	$\Omega = \{CC, CK, KC, KK\}$, com C = Cara e K = Coroa
c. Dois dados são lançados simultaneamente e estamos interessados na soma das faces observadas	$\Omega = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$

Evento é qualquer subconjunto do espaço amostral. Usualmente denotamos os eventos com as letras iniciais do alfabeto na forma maiúscula.

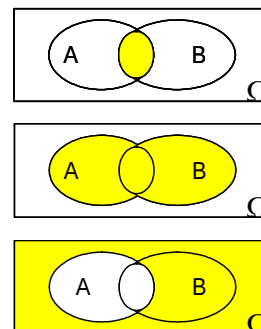
Exemplo 22. Considere o experimento de jogar um dado e observar o resultado. Alguns possíveis eventos desse experimento são: $A = \{\text{ocorrer a face 5}\} = \{5\}$ ou $B = \{\text{ocorrer face par}\} = \{2, 4, 6\}$ etc.

Existem dois eventos especiais: **espaço todo** (Ω) e o **conjunto vazio** ($\emptyset$). Esses eventos não têm aplicações práticas, mas serão úteis para provarmos propriedades das probabilidades.

Operações com eventos

Utilizando o diagrama de Venn, que foi introduzido em 1881 pelo filósofo e matemático britânico John Venn, podemos ilustrar as três operações básicas com eventos, a saber, interseção, união e complementar. Assim, sejam A e B dois eventos de um mesmo espaço amostral Ω .

- O evento interseção de A e B, denotado por $A \cap B$, é o evento em que A e B ocorrem simultaneamente.
- O evento união de A e B, denotado por $A \cup B$, é o evento em que A ocorre ou B ocorre (ou ambos).
- O evento complementar de A, denotado por A^c , é o evento em que A não ocorre.



Exemplo 23. Operações com eventos. Seja $\Omega = \{1, 2, 3, 4, 5, 6\}$. Considere os seguintes eventos: $A = \{2, 4, 6\}$, $B = \{4, 5, 6\}$ e $C = \{1, 3, 5\}$. Os eventos a seguir ficam assim:

$$A \cap B = \{4, 6\}$$

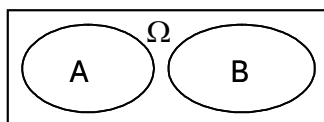
$$A \cap C = \emptyset$$

$$A \cup B = \{2, 4, 5, 6\}$$

$$A \cup B^c = \{1, 2, 3, 4, 6\}$$

Eventos disjuntos

Dois eventos A e B são **mutuamente exclusivos** ou **disjuntos** se eles não podem ocorrer simultaneamente ($A \cap B = \emptyset$).



Exemplo 24. Considere os seguintes eventos: $A = \{\text{o resultado do dado foi 4}\}$ e $B = \{\text{o resultado do dado foi 5}\}$. O evento $A \cap B = \emptyset$, pois é impossível existir o evento $A \cap B = \{\text{ocorrer 4 e 5, simultaneamente, em um único lançamento do dado}\}$.

Após essas quatro definições, acreditamos que o leitor esteja preparado para aprender a calcular probabilidades. Sugerimos assim, que faça os dois primeiros exercícios da seção 1.4 antes de prosseguir.

1.3 Definições de Probabilidade

A área de probabilidade começou a ser desenvolvida no século XVII antes ainda da formalização da área da Estatística, em questões propostas em jogos de azar. Em 1654, Pierre de Fermat (1601-1665) e Blaise Pascal (1623-1662), na França, estabelecem os **Princípios do Cálculo das Probabilidades**. Em 1656, Huygens (1629-1695) publica o primeiro **Tratado de Probabilidade**.



Fermat



Pascal



Huygens

No entanto, é fácil perceber que o termo **probabilidade** já está enraizado no senso comum, pois as pessoas vivem o cotidiano calculando **implicitamente** algumas probabilidades, tais como situações de sua vida pessoal; organizando-se em relação a horários a cumprir, levando em conta as circunstâncias do tráfego; agasalhando-se ao sair de casa se a previsão do tempo indicar uma frente fria. Em resumo, prevenindo-se em situações de risco.

A pergunta que surge então é “Como podemos definir Probabilidade?”.

Probabilidade é uma medida que quantifica a sua incerteza frente a um possível acontecimento futuro.

Há várias maneiras de se medir a incerteza e é costume se pensar na seguinte divisão:

1) Método Clássico

3) Método Subjetivo

2) Método Freqüentista

4) Método Moderno ou Axiomático

O primeiro é devido a Laplace e é o mais conhecido, pois relaciona eventos favoráveis com eventos possíveis. O segundo consiste em repetir um experimento várias vezes. O terceiro é baseado na opinião pessoal, e o último é devido a Kolmogorov e baseia-se no princípio de que qualquer experimento pode ser modelado.

Método Clássico

Consideremos o caso em que se joga um dado repetidas vezes. O dado tem seis faces: 1, 2, 3, 4, 5, 6. Se o dado é homogêneo, equilibrado, jogando-o uma vez não há razão para dizermos que determinada face tenha preferência sobre as outras. Todos os seis resultados são igualmente possíveis. Então a probabilidade de aparecer a face 3, por exemplo, é de $1/6$. O evento que nos interessa consiste em um elemento, e o espaço amostral tem seis elementos.

Definição 5.1. Se A é o evento de interesse, a probabilidade de A , representada por $P(A)$, é dada por

$$P(A) = \frac{\text{Número de casos favoráveis ao evento } A}{\text{Número de casos possíveis}} \quad (5.1)$$

Essa definição se aplica quando os pontos do espaço amostral são equiprováveis.

Exemplo 25. No lançamento de uma moeda equilibrada, qual a probabilidade de aparecer uma Cara? O espaço amostral associado é $\Omega = \{\text{Cara}, \text{Coroa}\}$. Pela definição clássica, a probabilidade de ocorrência do evento $A = \{\text{Cara}\}$ é $P(A) = 1/2$. Note que o número de elementos em Ω é 2 e o número de elementos em A é 1.

Método Freqüentista

A definição clássica de probabilidade só se aplica a espaços amostrais em que os eventos simples são igualmente possíveis. Esse é o caso da maioria das aplicações de probabilidades aos jogos de azar, área que, precisamente, suscitou os primeiros problemas práticos resolvidos pela teoria das probabilidades. Esses mesmos jogos, entretanto, repetidos inúmeras vezes, levaram a considerar a probabilidade de um evento como a freqüência relativa, ou seja, como a proporção de vezes que um evento ocorre em uma série suficientemente grande de realizações

de um experimento, em condições idênticas. Surgiu então uma nova definição de probabilidade, a definição freqüentista.

Definição 5.2. Se A é o evento de interesse, a probabilidade de A é dada por

$$P(A) = \frac{\text{Número de vezes que } A \text{ ocorreu}}{\text{Número total de repetições do experimento}} \quad (5.2)$$

em que o número de repetições deve ser grande.

Método Subjetivo

Definição 5.3. Cada indivíduo, baseado em informações anteriores e em sua opinião a respeito de um evento em questão, pode ter uma resposta para a probabilidade deste evento.

Exemplo 26. Um médico experiente consegue calcular uma probabilidade de o indivíduo ter uma determinada doença a partir dos sintomas que o indivíduo apresenta. Note que outro médico pode calcular uma probabilidade diferente para o mesmo indivíduo. Daí o caráter subjetivo.

Método Moderno

A definição clássica, freqüentista e subjetiva de probabilidade, embora sejam bastante intuitivas e devendo, por isso, ser sempre lembradas, não são definições matematicamente aceitáveis de probabilidade. Por exemplo, no caso da definição freqüentista, como saber se, à medida que o número de repetições de um experimento cresce, a freqüência relativa converge para um número. Além das dificuldades com o limite, existem muitas situações em que é necessário o uso de probabilidades, e, no entanto, não é nem possível nem intuitivo pensar em repetições.

A solução moderna consiste em axiomatizar algumas relações intuitivas e construir, a partir delas, toda a teoria de probabilidades, a exemplo do que se faz no estudo da geometria euclidiana.

Definição 5.4. Probabilidade é uma função $P(\cdot)$, que associa a cada evento do espaço amostral Ω , um número real, pertencente ao intervalo $[0, 1]$, satisfazendo os seguintes axiomas:

Axioma 1. $0 \leq P(A) \leq 1$.

Axioma 2. $P(\Omega) = 1$.

Axioma 3. Se A e B são eventos mutuamente exclusivos: $P(A \cup B) = P(A) + P(B)$.

A partir desses axiomas, podemos demonstrar as seguintes propriedades:

P1: $P(\emptyset) = 0$, onde $\emptyset$ é o conjunto vazio.

P2: Seja A^c o evento complementar de A, então $P(A^c) = 1 - P(A)$.

P3: Se A e B forem dois eventos quaisquer, então $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

P4: Se $A \subset B$, então $P(A) \leq P(B)$.

Exemplo 27. Seguem alguns exemplos de funções já descobertas na literatura para calcular probabilidades, que serão discutidas em detalhes nas próximas seções.

Distribuição	Função de probabilidades
Bernoulli	$P(X = x) = p^x (1 - p)^{1-x}$, $x = 0, 1$
Binomial	$P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}$, $x = 0, 1, \dots, n$
Hipergeométrica	$P(X = x) = \frac{\binom{r}{x} \binom{N-r}{n-x}}{\binom{N}{n}}$, $0 \leq x \leq \text{mínimo}(r, n)$.
Poisson	$P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}$, $x = 0, 1, \dots$
Uniforme	$f(x) = \frac{1}{\beta - \alpha}$, $\alpha < x < \beta$
Normal	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$, $-\infty < x < +\infty$

1.4 Exercícios

1. Determine o espaço amostral dos seguintes experimentos:

- a. Lançar 2 dados e observar as faces superiores;
- b. Lançar 2 dados e observar a soma das faces superiores;
- c. Uma urna contém 10 bolas azuis e 10 brancas. 3 bolas são retiradas ao acaso e as cores são anotadas;
- d. Uma moeda é lançada consecutivamente até o aparecimento da 1ª cara;
- e. Uma máquina produz 20 peças por hora. Ao final da primeira hora de produção, observa-se o nº de defeituosas;
- f. Medição do “tempo de vida” de uma lâmpada antes de se queimar:

2. Considere o seguinte espaço amostral: $\Omega = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$. Defina os eventos:

A = número par:

B = número ímpar:

C = múltiplo de 3:

D = maior ou igual a 6:

E = maior que 8:

F = menor que 5:

G = menor ou igual a 3:

Obtenha os seguintes eventos:

a. $A \cap B =$

e. $C \cap D =$

b. $A \cup B =$

f. $E \cup F =$

c. $(A \cap B)^c =$

g. $(A \cap G)^c =$

d. $(A \cup B)^c =$

h. $(E^c \cup B)^c =$

3. **Atividade Prática** do lançamento da moeda.

Passo 1 – Arrume um parceiro e tomem uma moeda – chamem o valor numérico da moeda de COROA (K) e a outra face de CARA (C). Suponham que haja interesse em saber se a sua moeda é “honesta” (isto significa saber se a probabilidade de CARA de sua moeda é 1/2 ou, em termos percentuais, se a chance de sair Cara é 50%).

Passo 2 – Um membro do grupo vai lançar a moeda e o outro vai marcar os resultados na planilha anexa, seguindo as seguintes instruções:

- Jogar a moeda uma vez e anotar C ou K no espaço adequado (linha 2) da planilha.
- Repetir este procedimento 30 vezes, preenchendo um a um todos os espaços da linha 2.

Passo 3 – Continuando com a planilha, trocar de lugar com o parceiro, voltar para os itens a) e b) das instruções e continuar mais 30 jogadas – até perfazer 60.

Passo 4 – Voltar ao primeiro da dupla e, ainda com a planilha, seguir as instruções:

- Depois do registro na linha 2 de todos os resultados como C ou K, passar para a linha 3: chamar CARA de 1 e COROA de 0 e colocar estes valores na planilha, abaixo de cada resultado já obtido na linha 2. Cada membro do grupo deve fazer metade – um faz a linha de cima e o outro a linha de baixo.

- Agora a linha 4 da planilha deve ser preenchida – em cada posição deve ser colocado o número acumulado de CARAS, até aquela jogada (verifique que a jogada está explicitada na linha 1- que é a linha **n**). Discutir com outro membro do grupo para ver se está claro – se não, pergunte! A linha de baixo é continuação do acumulado da linha de cima.

- Finalmente chegamos à última linha – linha 5: colocar a frequência relativa (m/n) de CARAS em cada momento – o que é isso? Discuta com o outro membro do grupo (**desprezar as entradas assinaladas com X**).

1) Jogada(n)	1	2	3	4	5	6	7	8	9	10	12	14		17		20			25				30
2) C ou K																							
3) 1 ou 0																							
4) Caras Acumuladas (m)																							
5) Frequência Relativa (m/n)											X	X	X	X	X	X	X	X	X	X	X	X	X

1) Jogada(n)	31	32	33							40						47		50				55				60
2) C ou K																										
3) 1 ou 0																										
4) Caras Acumuladas (m)																										
5) Frequência Relativa (m/n)	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X		X	X	X	X	X	X

Passo 5 – Depois de completar a 1ª parte da planilha, construir a seguinte tabela, usando as linhas 4 e 5 da planilha:

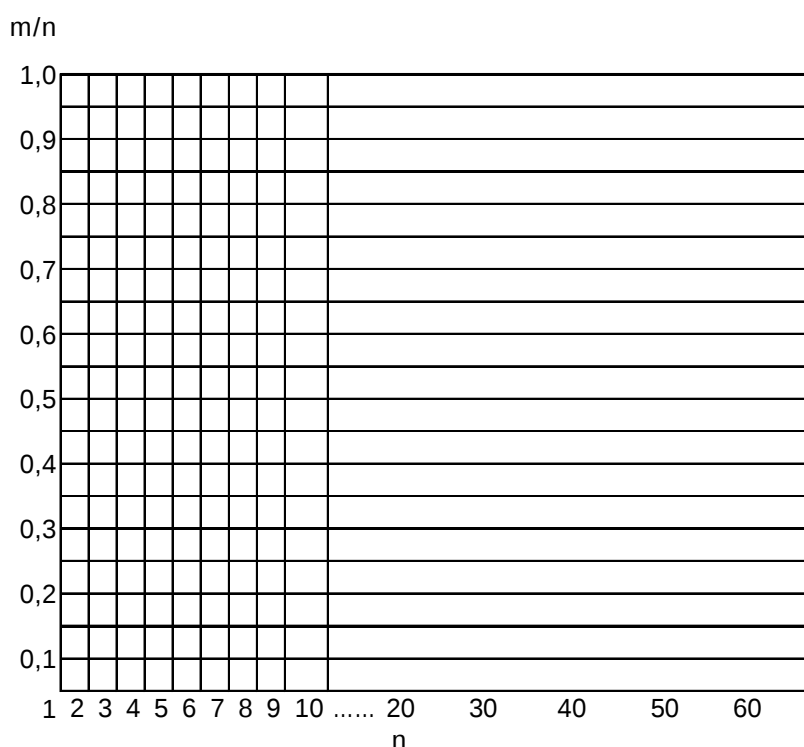
n	1	2	3	4	5	6	7	8	9	10	20	30	40	50	60
m/n															

Passo 6 – Completar o gráfico, usando os valores da tabela recém-construída, do seguinte modo:

Eixo Y – valores m/n Eixo X – valores da linha 1: (n)

Passo 7 – Comparar os resultados com os colegas e interpretar o resultado comentando sobre a “honestidade” da sua moeda.

Gráfico da Atividade Prática



Conclusão: com isso chegamos a uma possível “definição freqüentista” de probabilidade, ou seja, probabilidade é o valor em que a freqüência relativa se *estabiliza* após um número muito grande de ensaios.

Aula 2 – Fundamentos de Probabilidades

2.1 Probabilidade Condicional

A probabilidade condicional surge, por exemplo, quando se deseja calcular a probabilidade de um evento A ocorrer sabendo que um evento B já ocorreu.

Sejam A e B dois eventos associados a um mesmo espaço amostral Ω . Denota-se por $P(A|B)$ a probabilidade condicionada do evento A, quando o evento B tiver ocorrido.

Sempre que calculamos $P(A|B)$, estamos essencialmente calculando $P(A)$ em relação ao espaço amostral reduzido devido a B ter ocorrido, em lugar de fazê-lo em relação ao espaço amostral original Ω . Assim, uma definição mais formal de probabilidade condicional é dada pela definição 6.1.

Definição 6.1. Dados dois eventos A e B, a probabilidade condicional de A dado que ocorreu B é representada por $P(A | B)$ e definida por

$$P(A | B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) > 0 \quad (6.1)$$

Da expressão (6.1), obtemos a regra do produto de probabilidades dada por

$$P(A \cap B) = P(B)P(A | B) \quad (6.2)$$

Exemplo 28. Um grupo de pessoas foi classificado quanto a peso e pressão arterial, de acordo com as proporções do quadro a seguir:

Pressão	Peso			
	Excesso	Normal	Deficiente	Total
Alta	0,10	0,08	0,02	0,20
Normal	0,15	0,45	0,20	0,80
Total	0,25	0,53	0,22	1,00

- a. Qual a probabilidade de uma pessoa escolhida ao acaso nesse grupo ter pressão alta?
- b. Se se verifica que a pessoa escolhida tem excesso de peso, qual a probabilidade de ela ter também pressão alta?

Solução:

a. Como a pessoa é escolhida ao acaso em um grupo em que 20% tem pressão alta, chamando A o evento “ter pressão alta”, $P(A) = 0,20$ é a probabilidade pedida.

b. Chamemos B o evento “ter excesso de peso”. Nosso interesse passa a ser

$$P(A | B) = \frac{P(A \cap B)}{P(B)} = \frac{0,10}{0,25} = 0,40$$

O que fizemos foi precisamente estabelecer a probabilidade condicional de A dado B, $P(A|B)$, a partir de $P(A \cap B) = 0,10$ e $P(B) = 0,25$.

2.2 Independência de Eventos

Dois eventos A e B são independentes se a ocorrência de um não altera a probabilidade de ocorrência do outro, isto é, $P(A|B) = P(A)$ ou $P(B|A) = P(B)$, ou ainda, a seguinte forma equivalente:

$$P(A \cap B) = P(A) P(B) \quad (6.3)$$

Exemplo 29. Joaquina tem probabilidade de 0,8 de passar no vestibular, enquanto que Joãozinho tem probabilidade 0,6. Qual a probabilidade de ambos passarem no vestibular? Qual a suposição a ser feita nesse caso para calcular a probabilidade?

Solução: Sejam os eventos A: Joaquina passa no vestibular e B: Joãozinho passa no vestibular. Supondo independência entre os eventos A e B, temos que a probabilidade de ambos passarem no vestibular é $P(A \cap B) = 0,8 \times 0,6 = 0,48$.

2.3 Regra da Probabilidade Total

Considere a sequência $\{B_1, B_2, \dots, B_n\}$ como sendo uma partição do espaço amostral Ω , isto é, $B_i \cap B_j = \emptyset$ sempre que $i \neq j$ e $B_1 \cup B_2 \cup \dots \cup B_n = \Omega$. O diagrama da Figura 8 exibe uma partição de Ω .

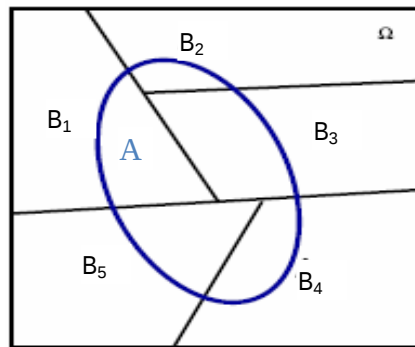


Figura 8. Partição do Ω e um evento qualquer A .

Vamos supor que o evento A possa ocorrer juntamente com um e só um dos n eventos mutuamente exclusivos $B_1, B_2, \dots, B_n$. Em outras palavras vamos assumir que

$$A = (A \cap B_1) \cup (A \cap B_2) \cup \dots \cup (A \cap B_n), \quad (6.4)$$

onde os eventos $A \cap B_i$ e $A \cap B_j$ (com subscritos distintos i e j) são mutuamente exclusivos.

Aplicando a função probabilidade em ambos os lados de (6.4) temos que

$$P(A) = P[(A \cap B_1) \cup (A \cap B_2) \cup \dots \cup (A \cap B_n)]. \quad (6.5)$$

Utilizando a regra de adição em (6.5) obtemos que

$$P(A) = P(A \cap B_1) + P(A \cap B_2) + \dots + P(A \cap B_n) \quad (6.6)$$

Através de (6.2), a expressão (6.6) fica

$$\begin{aligned} P(A) &= P(A | B_1) P(B_1) + P(A | B_2) P(B_2) + \dots + P(A | B_n) P(B_n) \\ &= \sum_{i=1}^n P(A | B_i) P(B_i) \end{aligned} \quad (6.7)$$

A expressão (6.7) é denominada **Regra da Probabilidade Total**.

Exemplo 30. Uma mineradora explora três minas denominadas B_1 , B_2 e B_3 . A partir de pesquisas anteriores, sabe-se que a probabilidade de encontrar ouro na mina B_1 é 0,1, na mina B_2 é 0,05 e na mina B_3 é 0,2. Além disso, essa mineradora tem explorado as minas B_1 , B_2 e B_3 nas proporções 0,3, 0,2 e 0,5, respectivamente. Qual a probabilidade de a mineradora encontrar ouro?

Solução: Seja $A = \{\text{encontrar ouro}\}$ e $B_j = \{\text{explorando a } j\text{-ésima mina } j\}$. Pela regra da probabilidade total temos

$$\begin{aligned} P(A) &= P(A | B_1) P(B_1) + P(A | B_2) P(B_2) + P(A | B_3) P(B_3) \\ &= 0,1 \times 0,3 + 0,05 \times 0,2 + 0,2 \times 0,5 = 0,14 \end{aligned}$$

2.4 Teorema de Bayes

Finalmente, uma das relações mais importantes envolvendo probabilidades condicionais é dada pelo teorema de Bayes. Thomas Bayes (1702-1761) afirmou que as probabilidades devem ser revistas quando conhecemos algo mais sobre os dados. A forma geral do teorema de Bayes pode ser introduzida através do Teorema 6.1.

Teorema 6.1. A probabilidade de ocorrência do evento B_i , supondo a ocorrência do evento A , é dado por

$$P(B_i | A) = \frac{P(A | B_i) P(B_i)}{\sum_{i=1}^n P(A | B_i) P(B_i)}$$

O teorema de Bayes é uma generalização da probabilidade condicional no caso de mais de dois eventos.

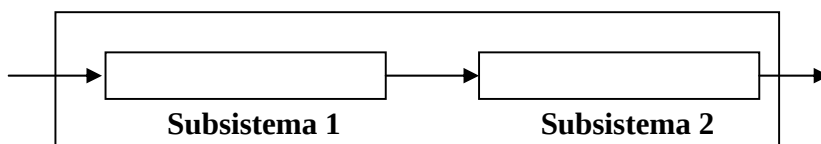
Exemplo 31. Considere novamente o Exemplo 30. Sabendo-se que a mineradora encontrou ouro, qual a probabilidade de que tenha sido na mina B_3 ?

Solução: Precisamos calcular a seguinte probabilidade:

$$P(B_3 | A) = \frac{P(A | B_3)P(B_3)}{\sum_{i=1}^n P(A | B_i)P(B_i)} = \frac{0,2 \times 0,5}{0,14} = \frac{0,10}{0,14} \cong 0,7143$$

2.5 Exercícios

1. O campo da Engenharia da confiabilidade se desenvolveu rapidamente a partir do início da década de 1960. Um tipo de problema encontrado é o de se estimar a confiabilidade de um sistema a partir das confiabilidades dos subsistemas. A confiabilidade é definida aqui como a probabilidade do funcionamento apropriado durante um certo período de tempo. Considere a estrutura de um sistema em série simples, como o da figura a seguir:



O sistema funciona se, e somente se, o subsistema 1 e o subsistema 2 funcionarem. Se os subsistemas sobrevivem independentemente, a confiabilidade do subsistema 1 é de 0,90 e do subsistema 2 é de 0,80, qual é a confiabilidade do sistema?

2. Em um centro de máquinas, há quatro máquinas automáticas de parafusos. Uma análise dos registros de inspeção passados fornece os seguintes dados:

Máquina	Percentual de Produção	Percentual de Defeituosos Produzidos
1	15	4
2	30	3
3	20	5
4	35	2

As máquinas 2 e 4 são mais novas e, assim, a maior parte da produção foi atribuída a elas. Suponha que o estoque atual reflita as porcentagens de produção indicadas.

- Se um parafuso é selecionado aleatoriamente do estoque, qual é a probabilidade de que seja defeituoso?
- Se um parafuso é selecionado aleatoriamente do estoque e ele é defeituoso, qual é a probabilidade de que seja da máquina 2?

UNIDADE IV

DISTRIBUIÇÕES DE PROBABILIDADES

Objetivos

Ao final desta unidade, você deverá ser capaz de

1. Associar variáveis aleatórias discretas com modelos probabilísticos.
2. Calcular probabilidades a partir do modelo Binomial, Hipergeométrico e de Poisson.
3. Associar variáveis aleatórias contínuas com modelos probabilísticos.
4. Calcular probabilidades a partir das distribuições Uniforme e Normal.

Aula 1 – Variáveis Aleatórias Discretas

1.1 Introdução

Vamos incorporar o conceito de probabilidade ao estudo de variáveis associadas a características em uma população. Muitos experimentos produzem resultados não-numéricos. Antes de analisá-los, é conveniente transformar seus resultados em números. Isso é feito através da *variável aleatória* (v.a.), que é uma função que associa um valor numérico a cada ponto do espaço amostral.

Para entender melhor o conceito, considere o exemplo que se segue.

Exemplo 32. Observa-se o sexo das crianças em famílias com três filhos. O espaço amostral é

$$\Omega = \{(MMM), (MMF), (MFM), (FMM), (MFF), (FMF), (FFM), (FFF)\}$$

Uma v.a. de interesse é $X = \{\text{nº. de crianças do sexo masculino}\}$. A cada evento simples ou ponto de Ω , associamos um número, que é o valor assumido pela v.a. X :

Evento	MMM	MMF	MFM	FMM	MFF	FMF	FFM	FFF
X	3	2	2	2	1	1	1	0

Poderíamos também ter considerado o número de crianças do sexo feminino. Os valores de X , na mesma ordem, seriam então 0, 1, 1, 1, 2, 2, 2, 3.

O passo fundamental para entendermos uma v.a. é associar a cada valor a sua probabilidade, obtendo assim a sua **distribuição de probabilidade**.

X	x_1	x_2	...	x_n
$P(X=x)$	$P(X=x_1)$	$P(X=x_2)$	...	$P(X=x_n)$

A função de probabilidade ($P(\cdot)$) deve satisfazer: $0 \leq P(X=x_i) \leq 1$ p/ $\forall x_i$ e $\sum_{i=1}^n P(X = x_i) = 1$.

Exemplo 33. Um certo departamento da UFSJ é formado por 35 professores, sendo 21 homens e 14 mulheres. Uma comissão de 3 professores será constituída, sorteando-se, ao acaso, três membros do departamento. Qual a probabilidade de a comissão ser formada pelo menos com duas mulheres?

Solução: Seja $X = \{\text{número de mulheres na comissão}\}$.

Espaço Amostral	X	Probabilidade
HHH	0	$\frac{21}{35} \times \frac{20}{34} \times \frac{19}{33} = 0,203$
HHM	1	$\frac{21}{35} \times \frac{20}{34} \times \frac{14}{33} = 0,150$
HMH	1	0,150
MHH	1	0,150
HMM	2	$\frac{21}{35} \times \frac{14}{34} \times \frac{13}{33} = 0,097$
MHM	2	0,097
MMH	2	0,097
MMM	3	$\frac{14}{35} \times \frac{13}{34} \times \frac{12}{33} = 0,056$

Distribuição de Probabilidade

X	0	1	2	3
P(X)	0,203	0,450	0,291	0,056

Assim, $P(X \geq 2) = P(X = 2) + P(X = 3)$

$$= 0,291 + 0,056$$

$$= \mathbf{0,347}$$

1.2 Esperança Matemática e Variância

Assim como definimos a média de uma distribuição de freqüências como a soma dos produtos dos diversos valores observados pelas respectivas freqüências relativas, é natural definirmos agora a média de uma v.a., ou de sua distribuição de probabilidade, como a soma dos produtos dos diversos valores de x_i da v.a. pelas respectivas probabilidades $P(X = x_i)$.

A média de uma v.a. X é também chamada valor esperado ou **esperança matemática**, ou simplesmente esperança de X . É representada por $E(X)$ e se define como

$$E(X) = x_1P(X = x_1) + x_2P(X = x_2) + \cdots + x_nP(X = x_n) = \sum_{i=1}^n x_iP(X = x_i)$$

É uma média ponderada dos x_i , em que os pesos são as probabilidades associadas.

Exemplo 34. Um lojista mantém extensos registros das vendas diárias de certo aparelho. O quadro a seguir dá o número x_i de aparelhos vendidos em uma semana e a respectiva probabilidade:

Número x_i	0	1	2	3	4	5
Probabilidade $P(X = x_i)$	0,1	0,1	0,2	0,3	0,2	0,1

Se for de R\$ 20,00 o lucro por unidade vendida, qual o lucro esperado nas vendas de uma semana?

Solução: Calculemos inicialmente $E(X)$, que é o número esperado de aparelhos vendidos em uma semana:

$$E(X) = (0)(0,1) + (1)(0,1) + (2)(0,2) + (3)(0,3) + (4)(0,2) + (5)(0,1) = 2,70.$$

Para x unidades vendidas o lucro é $20x$. Logo, o lucro esperado é de R\$ 54,00.

Variância

Assim como a média é uma medida de posição de uma v.a., é natural que procuremos uma medida de dispersão dessa variável em relação à média. Essa medida é a variância, a ser representada por σ^2 e definida por

$$\sigma^2 = \text{Var}(X) = E[X - E(X)]^2 = \sum_{i=1}^n (x_i - E(X))^2 P(X = x_i)$$

Desenvolvendo o termo quadrático do somatório, obtemos uma expressão mais fácil de calcular a variância dada por

$$\sigma^2 = \text{Var}(X) = E(X^2) - [E(X)]^2,$$

onde $E(X^2) = \sum_{i=1}^n x_i^2 P(X = x_i)$.

Desvio Padrão

O desvio padrão (σ) é a raiz quadrada positiva da variância. Tem sobre esse último a vantagem de exprimir a dispersão na mesma unidade de medida da v.a.

$$\sigma = \sqrt{\sigma^2}$$

1.3 Distribuições de Probabilidades para Variáveis Aleatórias Discretas

Para utilizar a teoria das probabilidades no estudo de um fenômeno concreto, devemos encontrar um modelo probabilístico adequado a tal fenômeno. Por modelo probabilístico para uma v.a. X entendemos uma forma específica de função de distribuição de probabilidade que reflita o comportamento de X .

Nesse processo de escolha, lançamos mão, em muitas situações, de algum modelo clássico. Nesta seção estudaremos os modelos discretos comumente utilizados: Bernoulli, Binomial, Hipergeométrica e Poisson.

1.3.1 Modelo Bernoulli

Na prática, existem muitos experimentos que admitem apenas dois resultados. Por exemplo, uma peça é classificada como boa ou defeituosa; um entrevistado concorda ou não com a afirmação feita; o resultado de um exame médico para detecção de uma doença é positivo ou negativo; no lançamento de um dado ocorre ou não a face 5.

Situações com alternativas **dicotômicas** podem ser representadas genericamente por respostas do tipo **sucesso-fracasso**. Esses experimentos recebem o nome de ensaio de **Bernoulli** e originam uma v.a. com distribuição Bernoulli.

Variável Aleatória de Bernoulli

É uma v.a. X que assume apenas dois valores: **1** se ocorrer **sucesso**, e **0** se ocorrer **fracasso**, sendo p a probabilidade de sucesso, $0 < p < 1$.

Denotamos por $X \sim \text{Bernoulli}(p)$ uma v.a. com distribuição de Bernoulli com parâmetro p , se

$$X = \begin{cases} 1, & \text{se ocorrer "sucesso"} \\ 0, & \text{se ocorrer "fracasso"} \end{cases} \quad \text{com função de probabilidade,}$$

$$P(X = x) = p^x (1-p)^{1-x}, \quad x = 0, 1$$

Daí segue que

$$E(X) = p \quad \text{e} \quad \text{Var}(X) = 1-p$$

Repetições independentes de um ensaio de Bernoulli dão origem ao modelo binomial.

1.3.2 Modelo Binomial

Um experimento é dito ser um experimento Binomial se

- a. consiste em n ensaios de Bernoulli;
- b. seus ensaios são independentes; e
- c. a probabilidade de sucesso em cada ensaio é sempre igual a p , $0 < p < 1$.

A v.a. X , correspondente ao número de sucessos num experimento binomial, tem distribuição binomial com parâmetros n e p , com função de probabilidade dada por

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \dots, n,$$

onde $\binom{n}{x} = \frac{n!}{x!(n-x)!}$, $n! = n(n-1)(n-2)\cdots(2)(1)$ e $0! = 1$.

Usamos a seguinte notação: $X \sim B(n; p)$. A média e a variância são dadas, respectivamente, por

$$E(X) = np \quad \text{e} \quad \text{Var}(X) = np(1-p)$$

Exemplo 35. Suponha que 20% dos clientes de uma empresa sejam inadimplentes. Se 10 pessoas dessa população forem escolhidas ao acaso e com reposição, determine

- a. O n° esperado de inadimplentes;
- b. A probabilidade de selecionar exatamente 3 pessoas inadimplentes;
- c. A probabilidade de selecionar no máximo 3 inadimplentes.

Solução:

a. $X = \{\text{número de pessoas inadimplentes}\}$. Temos que $E[X] = 10 \times 0,2 = 2$.

b. $P(X = 3) = \binom{10}{3} 0,2^3 (1 - 0,2)^{10-3} \cong 0,2$

c. $P(X \leq 3) = \sum_{i=0}^3 P(X = i) = 0,8^{10} + \binom{10}{1} 0,2^1 0,8^9 + \binom{10}{2} 0,2^2 0,8^8 + \binom{10}{3} 0,2^3 0,8^7 \cong 0,88$

1.3.3 Modelo Hipergeométrico

A distribuição hipergeométrica está restritamente relacionada com a distribuição binomial. A diferença-chave entre as duas distribuições de probabilidade é que, com a distribuição hipergeométrica, os ensaios não são independentes, e a probabilidade de sucesso muda de ensaio para ensaio, pois as seleções dos elementos são feitas sem reposição, enquanto que na distribuição binomial as seleções dos elementos são feitas com reposição.

Considere um conjunto de N objetos dos quais (r) são do tipo I e $(N - r)$ são do tipo II. Um sorteio de n objetos ($n < N$) é feito ao acaso e sem reposição. A variável aleatória discreta X que é igual ao número de objetos do tipo I selecionados nesse sorteio tem distribuição hipergeométrica.

Os valores possíveis de X vão de 0 a $\min(r, n)$, uma vez que não podemos ter mais do que o número de objetos existentes do tipo I, nem mais que o total de sorteados.

Sua função de probabilidade é dada por

$$P(X = x) = \frac{\binom{r}{x} \binom{N-r}{n-x}}{\binom{N}{n}}, \quad 0 \leq x \leq \min(r, n).$$

Usamos a seguinte notação: $X \sim \text{Hipergeométrica}(N; n; r)$. A esperança e variância são dadas por

$$E(X) = np \quad \text{e} \quad \text{Var}(X) = np(1-p)(N-n)/(N-1),$$

onde $p = r/N$.

Exemplo 36. Uma fábrica produz peças que são embaladas em caixas com 40 unidades. Para aceitar o lote de caixas enviado por essa fábrica, o controle de qualidade de uma empresa sorteia uma caixa do lote e sorteia 10 peças, sem reposição, dessa mesma caixa. Se houver alguma peça defeituosa, o lote inteiro é devolvido. Se a caixa sorteada tiver 4 peças defeituosas, qual é a probabilidade de o lote não ser devolvido?

Solução: $N = 40$, $n = 10$ e $r = 4$. X : número de peças defeituosas.

$$P(X = 0) = \frac{\binom{4}{0} \binom{40-4}{10-0}}{\binom{40}{10}} \cong 0,3$$

1.3.4 Modelo Poisson

A distribuição de Poisson é empregada em experimentos nos quais não se está interessado no número de sucessos obtido em n tentativas, como ocorre no caso da distribuição binomial, mas sim no número de sucessos ocorridos durante um intervalo contínuo, que pode ser um intervalo de tempo, espaço etc. Alguns exemplos de variáveis que podem ter a distribuição de Poisson são

- número de defeitos por centímetro quadrado;
- nº de acidentes por dia;

- n° de clientes por hora;
- n° de chamadas telefônicas recebidas por minuto.

Note-se que a unidade de medida (tempo, área) é **contínua**, mas a variável aleatória de interesse (número de ocorrência) é **discreta**. Além disso, as falhas não são contáveis. Não é possível contar os acidentes que não ocorreram, nem o número de defeitos por centímetros quadrados que não ocorreram.

O limite inferior do número de ocorrências, em todas as situações dos exemplos, é zero, enquanto que o limite superior é – ao menos teoricamente – infinito, muito embora, na maioria dos exemplos acima, seja difícil imaginar um número infinito de ocorrências.

As probabilidades, calculadas agora para todos os números inteiros não negativos $x = 0, 1, 2, \dots$ são dadas da seguinte forma:

$$P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x = 0, 1, \dots,$$

onde “ X = números de sucessos em um intervalo” é a variável de interesse, $\lambda > 0$ é o número médio de sucessos da variável X e “ e ” é a constante 2,7183 (base dos logaritmos naturais).

Usamos a seguinte notação: $X \sim P(\lambda)$. A esperança e variância são dadas por

$$E(X) = \text{Var}(X) = \lambda$$

Exemplo 37. Um departamento de conserto de máquinas recebe uma média de cinco chamadas por hora. Supondo que a distribuição de Poisson seja adequada nessa situação, obter a probabilidade de que, em uma hora selecionada aleatoriamente, sejam recebidas exatamente três chamadas.

Solução: Seja X : número de chamadas para conserto de máquinas em uma hora. O parâmetro $\lambda = 5/\text{hora}$. Aplicando na função da Poisson, temos

$$P(X = 3) = \frac{e^{-5} 5^3}{3!} \cong 0,14$$

1.4 Exercícios

1. A distribuição de X : nº de crianças por domicílio numa determinada região é dada pela tabela abaixo.

X	0	1	2	3	4	5
$P(X = x)$	0,10	0,15	0,25	0,30	0,15	0,05

Calcule:

- a. O número médio de crianças por domicílio, μ_X .
- b. O desvio padrão de X , σ_X .
- c. A probabilidade $P\{\mu_X - \sigma_X \leq X \leq \mu_X + \sigma_X\}$.

2. Sabe-se que 7% dos ratos machos de uma certa linhagem são portadores de um defeito genético que não ocorre em fêmeas. Responda:

- a. Qual a probabilidade de encontrarmos pelo menos 1 animal com esse defeito genético numa ninhada com 5 machos?
- b. Qual a probabilidade de encontrarmos no máximo 3 animais com esse defeito genético numa ninhada com 4 machos?

3. Numa central telefônica, o número de chamadas chega segundo uma distribuição Poisson, com a média de oito chamadas por minuto. Determine qual a probabilidade de que num minuto se tenha(m)

- a. duas ou mais chamadas;
- b. menos que duas chamadas;
- c. entre sete (inclusive) e nove (exclusive) chamadas.

Aula 2 – Variáveis Aleatórias Contínuas

2.1 Introdução

Até aqui estudamos variáveis aleatórias discretas que são caracterizadas por ter uma distribuição de probabilidade dada por uma tabela que associa a cada um de seus valores uma probabilidade. Esta probabilidade é um número entre 0 e 1 cuja soma é igual a 1. Vamos agora definir uma variável aleatória contínua.

Seja X uma variável aleatória. Suponha que os possíveis valores de X sejam um intervalo que possui infinitos valores; então, dizemos que X é uma **variável aleatória contínua**.

Exemplo 38. Seguem alguns exemplos de variáveis aleatórias contínuas.

- a. Mede-se a altura de uma mulher em uma cidade. O valor encontrado é um número real. Aqui também sabemos que esse número não passa de 3 metros, mas é conveniente considerar qualquer número real positivo.
- b. Em campanhas preventivas de hipertensão arterial é comum, de tempos em tempos, medir-se o nível de colesterol. O valor de cada medida pode ser um número real não-negativo.
- c. Retira-se uma lâmpada da linha de produção e coloca-se a mesma em um soquete, acendendo-a; observa-se a mesma até que se queime. O tempo de duração da lâmpada é um número real não negativo.

No Exemplo 38 o número observado em cada um dos experimentos aleatórios é um número real e resulta em geral de uma medição: altura das mulheres; nível de colesterol e tempo de duração da lâmpada.

Uma variável aleatória contínua assume seus possíveis valores em um determinado intervalo. A pergunta que surge é “Como são atribuídas probabilidades neste caso?”.

Exemplo 39. Suponha que observamos o *peso*, em kg, de 1500 pessoas adultas selecionadas aleatoriamente numa população. O histograma por densidade desses valores é apresentado na Figura 9.

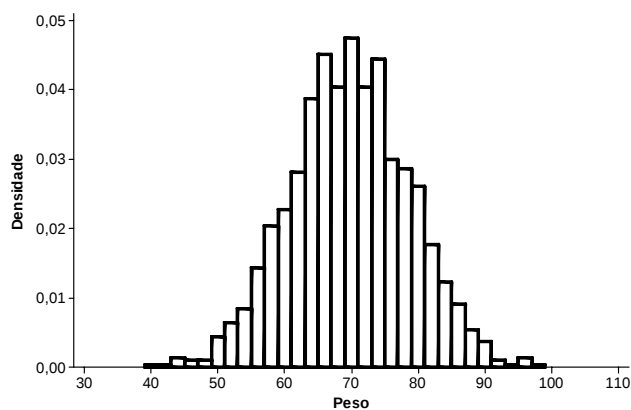


Figura 9. Histograma da variável *peso*.

A análise do histograma indica que a distribuição dos valores da variável *peso* é aproximadamente **simétrica** em torno de 70 kg; a maioria dos valores encontra-se no intervalo (50; 90); existe uma pequena proporção de valores abaixo de 50 kg e acima de 90 kg.

Seja $X = \{\text{peso em kg}\}$ de uma pessoa adulta escolhida ao acaso da população. Como se distribuem os valores da v.a. X , ou seja, qual a distribuição de probabilidades de X ?

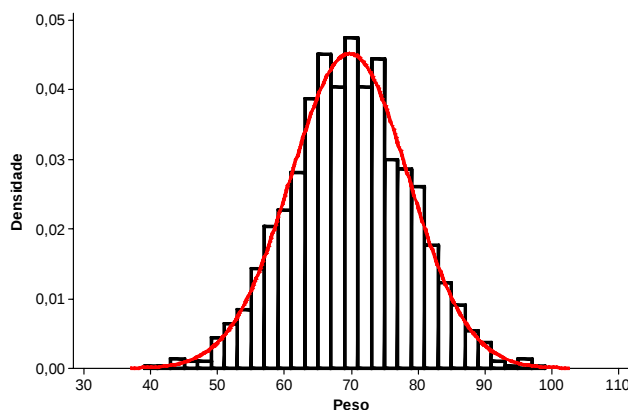


Figura 10. Histograma da variável *peso* com o ajuste da distribuição normal.

A Figura 10 ilustra o histograma da variável *peso* apresentado na Figura 9 com o ajuste de uma função densidade, conhecida como distribuição normal.

Para as variáveis contínuas, as probabilidades são atribuídas por meio de uma função cuja área entre a função e o eixo das abscissas (X) é igual a um.

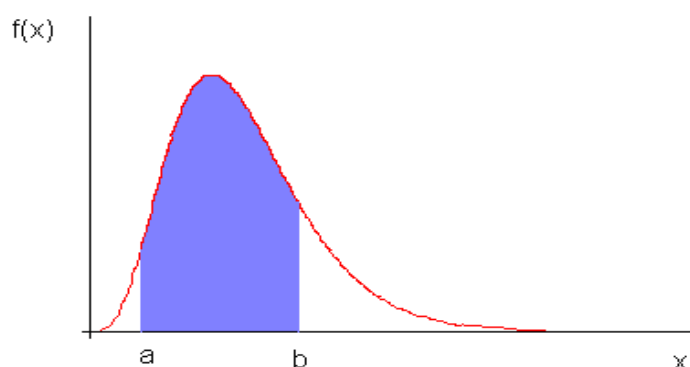


Figura 11. Representação de uma função densidade de probabilidade contínua.

A área hachurada na Figura 11 ilustra a probabilidade de a v.a. contínua X estar no intervalo $[a, b]$, ou seja, $P(a \leq X \leq b) = \text{área hachurada}$.

Esta função $f(x)$ é denominada função densidade de probabilidade (fdp) da variável aleatória contínua X . A área sob uma curva delimitada por dois valores a e b , como mostra a Figura 11 é determinada calculando-se a integral definida entre a e b da densidade de probabilidade representada pela função, isto é,

$$\int_a^b f(x)dx = P(a \leq x \leq b)$$

Exemplo 40. Um fabricante de televisão a cores oferece uma garantia de 1 ano para substituição gratuita se o tubo de imagem falhar. Ele estima o tempo de falha (em unidades de anos), x , como uma variável aleatória contínua com a seguinte fdp

$$f(x) = \begin{cases} \frac{1}{4}e^{-x/4}, & x > 0 \\ 0 & x \leq 0 \end{cases}.$$

Qual a probabilidade de você comprar a televisão e necessitar de uma substituição gratuita?

Solução:

$$P(x \leq 1) = \int_0^1 \frac{1}{4} e^{-\frac{x}{4}} dx \cong 0,2$$

Função Densidade de Probabilidade

Se X é uma v.a. contínua, a **função densidade de probabilidade** $f(X)$, indicada abreviadamente por fdp, é uma função que satisfaz às seguintes condições:

a. $f(X) \geq 0, \forall X$;

b. A área sob a função densidade de probabilidade é 1, isto é: $\int_{-\infty}^{+\infty} f(x)dx = 1$;

c. $P(a \leq X \leq b)$ = área sob a função densidade de probabilidade $f(x)$ e acima do eixo x entre os pontos a e b , isto é, $P(a \leq x \leq b) = \int_a^b f(x)dx$;

d. $P(X = x_0) = 0$, porque, $P(X = x_0) = \int_{x_0}^{x_0} f(x)dx = 0$. Como consequência, temos

$$P(a < X < b) = P(a \leq X < b) = P(a < X \leq b) = P(a \leq X \leq b).$$

Função de Distribuição Acumulada

Se X é uma v.a. contínua, a **função de distribuição acumulada** (fda) de X é definida como

$$F(X) = P(X \leq x) = \int_{-\infty}^x f(s)ds.$$

Exemplo 41. Considere a seguinte densidade de probabilidade: $f(x) = 2x$, para $0 \leq x \leq 1$ e $f(x) = 0$, fora desse intervalo. Obtenha a $F(x)$ de X .

Solução:

$$F(x) = \begin{cases} 0, & x < 0 \\ \int_0^x 2sds = s^2 \Big|_0^x = x^2, & 0 \leq x \leq 1 \\ 1 & x > 1 \end{cases}$$

2.2 Esperança Matemática e Variância

Se X é uma v. a. contínua, o **valor esperado de X** (ou **esperança matemática** de X) denotada por $E(X)$ é definido como

$$E[X] = \int_{-\infty}^{+\infty} xf(x)dx$$

Exemplo 42. Para uma variável que tem densidade $f(x) = 2x$, $0 < x < 1$, então,

$$E[X] = \int_0^1 x 2x dx = \int_0^1 2x^2 dx = \frac{2}{3}x^3 \Big|_0^1 = \frac{2}{3}.$$

A variância de uma variável aleatória contínua é definida por:

$$\text{Var}(X) = E(X^2) - [E(X)]^2, \text{ onde } E[X^2] = \int_0^1 x^2 f(x) dx.$$

Exemplo 43. Para uma variável que tem densidade $f(x) = 2x$, $0 < x < 1$, calcule a variância de X , sabendo que $E[X] = \frac{2}{3}$ do Exemplo 42.

Solução: $E[X^2] = \int_0^1 x^2 2x dx = \int_0^1 2x^3 dx = \frac{2}{4}x^4 \Big|_0^1 = \frac{2}{4}$. Logo, $\text{Var}[X] = 2/4 - (2/3)^2 = 1/18$.

Conseqüentemente, o desvio padrão de X fica $DP[X] = \sqrt{\text{Var}[X]} = \sqrt{1/18} \cong 0,236$

2.3 Distribuições de Probabilidades para Variáveis Aleatórias Contínuas

As distribuições discretas de probabilidades tratam de situações em que o espaço amostral contém um número finito, ou infinito enumerável, de pontos. Se o espaço amostral contém um número infinito não-enumerável de pontos, temos que trabalhar com as distribuições contínuas

de probabilidades. Abordaremos aqui, em caráter mais intuitivo, a distribuição uniforme e a distribuição normal.

2.3.1 Modelo Uniforme

A distribuição de probabilidade mais simples de uma v.a. X contínua é a distribuição uniforme.

Uma v.a. X tem distribuição uniforme $U(a, b)$ se sua função densidade de probabilidade é da forma

$$f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{caso contrário} \end{cases}.$$

A média e a variância da distribuição $U(a, b)$ são dadas respectivamente por

$$E[X] = \frac{a+b}{2} \quad \text{e} \quad \text{Var}[X] = \frac{(b-a)^2}{12}$$

Note que a média é exatamente o ponto médio do intervalo $[a, b]$.

Exemplo 44. Devido à presença de quantidades variáveis de impureza, o ponto de fusão de certa substância pode ser considerado uma v.a. contínua distribuída uniformemente no intervalo $[100, 125]$. Qual a probabilidade de a substância fundir-se entre 110 e 115?

Solução: Neste caso, $a = 100$, $b = 125$ e $b - a = 25$. A função densidade fica

$$f(x) = \begin{cases} \frac{1}{25}, & 100 \leq x \leq 125 \\ 0, & \text{caso contrário} \end{cases}$$

A probabilidade procurada é

$$P(110 < X < 115) = \int_{110}^{115} \frac{1}{25} dx = \frac{1}{25} x \Big|_{110}^{115} = \frac{115-110}{25} = \frac{5}{25} = 0,2$$

2.3.2 Modelo Normal

A distribuição Normal é a mais importante das distribuições contínuas de probabilidade. Foi introduzida em 1730 por D'Moivre, sendo muito utilizada em Astronomia pelo alemão físico e matemático Gauss, trazendo muita confusão para várias pessoas que, por esse motivo, acham que foi Gauss quem a descobriu. Muitos dos fenômenos aleatórios de interesse comportam-se próximos a essa distribuição com valores muito freqüentes em torno da média e diminuindo a freqüência à medida que nos afastamos da média.

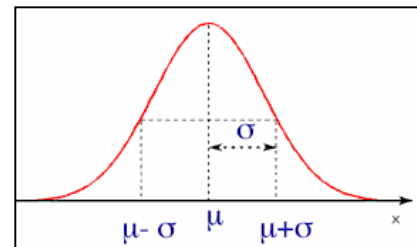
A distribuição normal tem sua densidade dada por

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad -\infty < x < \infty$$

em que μ e σ são os parâmetros da distribuição.

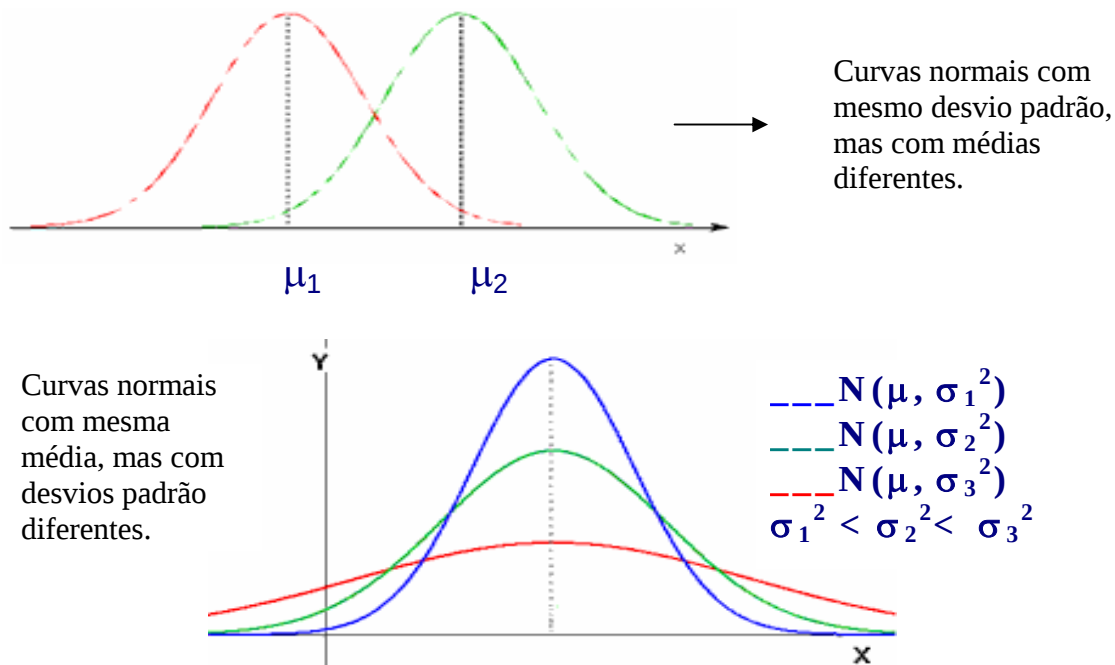
As principais características da distribuição normal são:

- A média da distribuição é μ ;
- O desvio padrão é σ ;
- A moda e a mediana são iguais a μ ;
- A curva normal é simétrica em torno da média μ ;
- Os pontos de inflexão são $\mu - \sigma$ e $\mu + \sigma$;
- A área sob a curva e acima do eixo horizontal é igual a 1.



A v.a. Normal com média μ e variância σ^2 é denotada por $N(\mu, \sigma^2)$.

A distribuição normal depende dos parâmetros μ e σ^2



A variável Normal Padronizada

O cálculo direto de probabilidades envolvendo a distribuição normal exige recursos de cálculo infinitesimal e, mesmo assim, dada a forma da função de densidade, não é um processo elementar. Por isso, elas foram tabeladas, permitindo-nos obter diretamente o valor da probabilidade desejada.

Notemos, entretanto, que a função de densidade normal depende de dois parâmetros, μ e σ , de modo que, se as probabilidades fossem tabeladas diretamente a partir dessa função, seriam necessárias tabelas de dupla entrada, complicando-se consideravelmente. Recorre-se, por isso, a uma mudança de variável, transformando a v.a. X na v.a. Z assim definida:

$$Z = \frac{X - \mu}{\sigma}.$$

Essa nova variável chama-se variável normal padronizada. Recebe esse nome, porque sua média é 0 e seu desvio padrão é 1. Mediante tal transformação, basta construirmos uma única

tabela, a da normal reduzida e, através dela, obteremos as probabilidades associadas a todas as distribuições $N(\mu, \sigma^2)$.

Note que essa transformação não altera a forma da distribuição, apenas refere-se a uma nova escala.

Assim, se quisermos calcular $P(a < X < b)$, sendo $X \sim N(\mu; \sigma^2)$, podemos definir $Z = \frac{X - \mu}{\sigma}$ e calcular a seguinte probabilidade:

$$P(a < X < b) = P(a - \mu < X - \mu < b - \mu) = P\left(\frac{a - \mu}{\sigma} < \frac{X - \mu}{\sigma} < \frac{b - \mu}{\sigma}\right) = P\left(\frac{a - \mu}{\sigma} < Z < \frac{b - \mu}{\sigma}\right)$$

Uma representação do cálculo dessa probabilidade é apresentada na Figura 12.

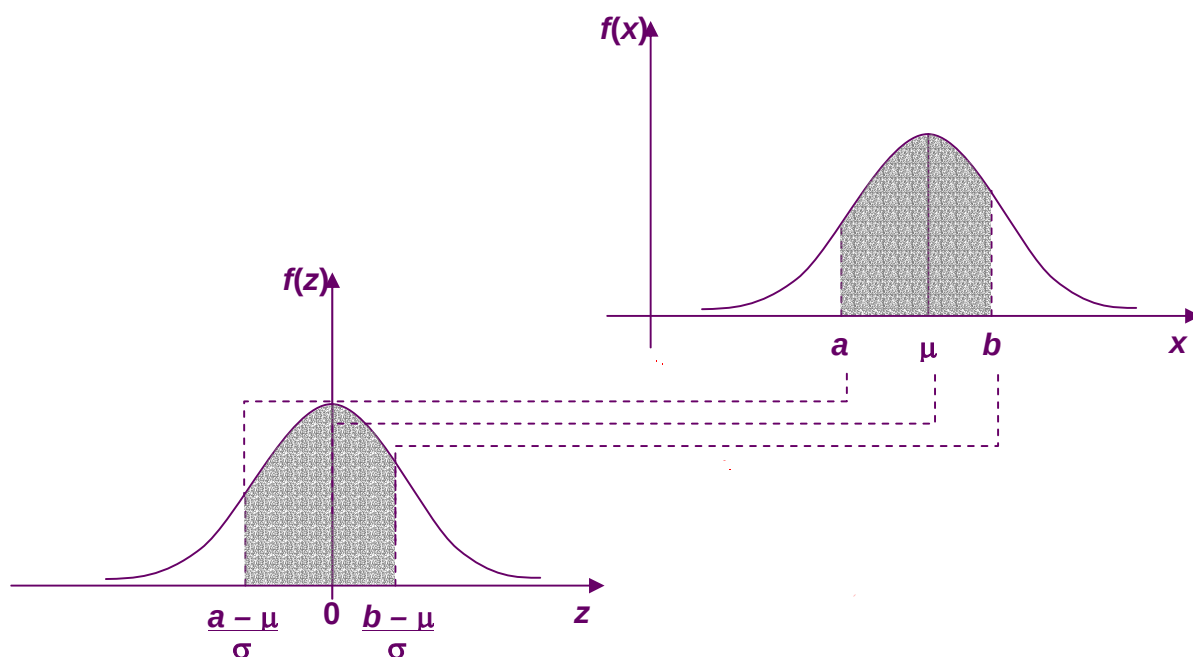


Figura 12. Representação do cálculo da $P(a < X < b)$ via variável normal padronizada Z .

De forma análoga, dada uma variável padronizada $Z \sim N(0;1)$, podemos obter a v.a. $X \sim N(\mu, \sigma^2)$ através da transformação inversa $X = \mu + Z\sigma$.

Tabela da Distribuição Normal Padrão

Denotamos: $A(z) = P(Z \leq z)$, para $z \geq 0$.



Probabilidades Acumuladas da Distribuição Normal (0, 1)

$A(z) = P(Z \leq z)$, $z \geq 0$.

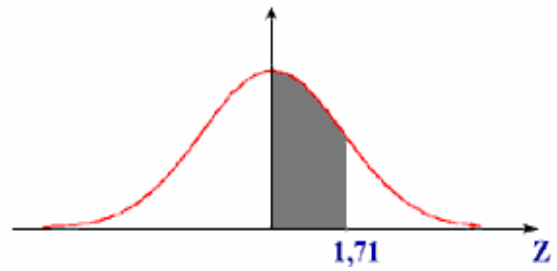
		Segunda decimal de z									
		0	1	2	3	4	5	6	7	8	9
Parte inteira e primeira decimal de z	0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
	0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
	0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
	0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
	0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
	0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
	0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
	0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
	0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
	0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
	1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
	1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
	1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
	1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
	1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
	1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
	1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
	1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
	1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
	1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
	2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
	2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
	2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
	2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
	2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
	2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
	2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
	2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
	2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
	2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
	3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
	3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
	3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
	3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
	3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
	3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
	3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
	3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
	3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
	3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Exemplo 45. Se $Z \sim N(0,1)$, então:

a. $P(Z \leq 1,71) = A(1,71) = 0,9564$



b. $P(0 < Z \leq 1,71) = A(1,71) - A(0)$
 $= 0,9564 - 0,5000 = 0,4564$



Exemplo 46. Seja $X = \{\text{gasto com lanche semanal}\}$. Após estudar esta variável, vimos que $X \sim N(20, 64)$, então obtenha

a. $P(16 < X < 22)$

Solução:

$$P(16 < X < 22) = P\left(\frac{16-20}{8} < \frac{X-20}{8} < \frac{22-20}{8}\right) = P(-0,5 < Z < 0,25)$$

$$= (A(0,25) - A(0)) + (A(0,5) - A(0)) = (0,5987 - 0,5) + (0,6915 - 0,5) = 0,2902$$

b. $P(X < 18 \text{ ou } X > 24)$

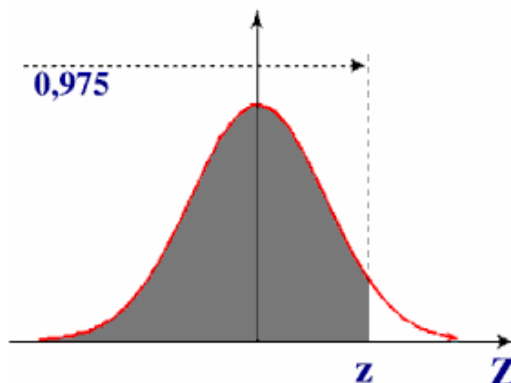
Solução:

$$P(X < 18 \text{ ou } X > 24) = P(X < 18) + P(X > 24) = P\left(\frac{X-20}{8} < \frac{18-20}{8}\right) + P\left(\frac{X-20}{8} > \frac{24-20}{8}\right)$$

$$= P(Z < -0,25) + P(Z > 0,5) = (1 - A(0,25)) + (1 - A(0,5))$$

$$= (1 - 0,5987) + (1 - 0,6915) = 0,7098$$

Como encontrar o valor z da distribuição $N(0,1)$ tal que área acumulada até ele seja $A(z) = 0.975$.



a. $P(Z \leq z) = 0,975$. z é tal que $A(z) = 0,975$. Pela tabela, $z = 1,96$.

Considere que $X \sim N(\mu, \sigma^2)$. Calcule k tal que $P(X \geq k) = 0,05$. Neste caso temos que

$$P(X \geq k) = P\left(\frac{X - \mu}{\sigma} \geq \frac{k - \mu}{\sigma}\right) = P\left(Z \geq \frac{k - \mu}{\sigma}\right) = 0,05 \Rightarrow A\left(\frac{k - \mu}{\sigma}\right) = 0,95$$

$$\Rightarrow \frac{k - \mu}{\sigma} = 1,64 \Rightarrow k = \mu + 1,64 \sigma$$

Logo, o valor de k é **$k = \mu + 1,64 \sigma$** .

Nota Importante: Para toda v.a. $X \sim N(\mu ; \sigma^2)$ temos

1. $P(\mu - \sigma \leq X \leq \mu + \sigma) = P(-1 \leq Z \leq 1) = 0,683$.
2. $P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = P(-2 \leq Z \leq 2) = 0,955$
3. $P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) = P(-3 \leq Z \leq 3) = 0,997$

2.4 Exercícios

1. Se $Z \sim N(0,1)$, calcule:

- | | |
|-----------------------------------|-------------------------------------|
| a. $P(1,32 < Z \leq 1,79)$ | d. $P(Z \geq 1,5)$ |
| b. $P(Z \leq -1,3)$ | e. $P(-1,5 \leq Z \leq 1,5)$ |
| c. $P(-1,32 < Z < 0)$ | f. $P(-2,3 < Z \leq -1,49)$ |

2. Encontre o valor z da distribuição $N(0,1)$ tal que

- | | |
|--|--------------------------------|
| a. $P(0 < Z \leq z) = 0,4975$ | d. $P(Z \geq z) = 0,3$ |
| b. $P(Z \geq z) = 0,975$ | e. $P(Z \leq z) = 0,10$ |
| c. $P(-z \leq Z \leq z) = 0,80$ | |

3. O diâmetro de um cabo elétrico é uma variável aleatória com fdp dada por

$$f(x) = \begin{cases} 6x(1-x), & 0 < x < 1 \\ 0, & \text{caso contrário} \end{cases}$$

- a.** Verifique se $f(x)$ é uma fdp, através do item **b.** da definição 8.2.
b. Obtenha a $F(x)$.
c. Qual a probabilidade de o diâmetro ser:

- | | |
|----------------------------|-------------------------------|
| c1. Igual a 0,5 cm? | c3. Entre 0,10 e 0,20? |
| c2. Maior que 0,5? | c4. Menor que 1? |

4. A quantia gasta anualmente, em milhões de reais, na manutenção do asfalto de uma cidade do interior é representada pela variável y , modelada pela função

$$f(x) = \begin{cases} \frac{8y-4}{9}, & 0,5 \leq y \leq 2 \\ 0, & \text{caso contrário} \end{cases}.$$

Qual a probabilidade de a quantia gasta ser inferior a 0,8 milhões de reais?

5. O tempo de sobrevivência de uma bateria (em anos) pode ser modelado pela função

$$f(x) = \begin{cases} e^{-x}, & x \geq 0 \\ 0, & \text{caso contrário} \end{cases}$$

a. Qual a probabilidade de a bateria sobreviver mais que 2 anos?

b. Qual é o tempo médio de sobrevivência da bateria?

6. O tempo gasto no exame vestibular de uma universidade tem distribuição Normal, com $\mu = 120$ min, e $\sigma = 15$ min.

a. Sorteando-se um aluno ao acaso, qual é a probabilidade de ele terminar o exame antes de 100 minutos?

b. Qual deve ser o tempo de prova, de modo a permitir que 95% dos vestibulandos a terminem no prazo estipulado?

c. Qual o intervalo central de tempo tal que 80% dos estudantes gaste para completar o exame?

REFERÊNCIAS

- BUSSAB, W. O.; MORETTIN, P. A. **Estatística Básica**. 5. ed. São Paulo: Saraiva, 2006.
- FARIAS, A. A.; SOARES, J. F.; CÉSAR, C. C. **Introdução à Estatística**. 2. ed. Rio de Janeiro: LTC, 2003.
- GNEDENKO, B. V. **A teoria da probabilidade**. Rio de Janeiro: Ciência Moderna, 2008.
- JAMES, B. R. **Probabilidade: um curso em nível intermediário**. 2. ed. Rio de Janeiro: SBM, 1996.
- MAGALHÃES, M. N.; PEDROSO DE LIMA, A. C. **Noções de Probabilidade e Estatística**. 6. ed. São Paulo: Edusp, 2007.
- ROSS, Sheldon. **A first course in probability**. 8. ed. Londres: Prentice Hall, 2005.
- TRIOLA, M. F. **Introdução à estatística**. 10. ed. Rio de Janeiro: LTC, 2008.